



Pyramid-KNet: Pyramid Kroneckernet-based lung cancer detection using computed tomography images

Praveen Kantha¹ · J. Anitha² · Katakam Venkateswara Rao³ · Balajee Maram⁴ · Satish Muppudi⁵ · Parul Datta⁶

Received: 8 July 2024 / Accepted: 15 February 2025 / Published online: 25 March 2025

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2025

Abstract

Lung Cancer (LC) is one of the most significant life-threatening cancers globally. Early detection and treatment are essential for patient recovery. Clinical professionals use histopathological images of biopsied lung tissue for diagnosis. However, identifying LC types is often error-prone and time-consuming. To address the above-mentioned challenges, an effective model Pyramid-KNet is developed for detecting LC using Computed Tomography (CT) images. Initially, the CT images are subjected to image enhancement using Histogram equalization. After that, lung lobe segmentation is completed by employing Psi-Net. Then, the lung nodule identification is conducted by the grid-based strategy. Furthermore, feature extraction is conducted for extracting features like Weber local descriptor (WLD), Median Robust Extended Local Binary Pattern with Discrete Cosine Transform (MRELBP with DCT), Texton, Gray-Level Co-occurrence Matrix (GLCM), and statistical features. Finally, LC detection is executed by employing hybrid Pyramid KroneckerNet (Pyramid-KNet), which is the integration of PyramidNet and Deep Kronecker Network (DKN), where layers are modified employing the Taylor concept. Furthermore, the performance of Pyramid-KNet is validated by comparing it with the performance of baseline methods, and Pyramid-KNet attained superior performance with an accuracy of 93%, precision of 92% and F-measure of 94% respectively.

Keywords Lung cancer (LC) · Computed tomography (CT) · Gray-Level Co-occurrence Matrix (GLCM) · Weber local descriptor (WLD) · PyramidNet

1 Introduction

Cancer is a common and serious disease characterized by the growth and spread of abnormal cells within the body. Early diagnosis significantly enhances the effectiveness of treatment. [1]. The abnormal cells can spread into other body organs [2, 3]. LC is the foremost reason for cancer deaths globally. The two common forms of LC are non-smaller cell LC and smaller cell LC [4]. Even though, the LC is detected earlier and cured with regular treatment, non-smaller cells are frequently diagnosed at an advanced stage [5]. LC is the type of cancer that affects the lung tissue [3, 6]. Also, LC refers to the deadly disease globally and it speedily spreads to other portions of the body like lymph nodes and even in the brain. Presently, on the basis of the estimation of the World Health Organization (WHO), about 7.8 million citizens die globally in a year due to LC. Moreover, it states that, if the condition of cancer is increasing continuously, it may affect more than 18 million people in the upcoming years [7]. With appropriate diagnosis and proper treatment, a possibility that one can cure LC. Thus, LC screening [34]

✉ Satish Muppudi
satish.m@gmrit.edu.in

¹ School of Engineering & Technology, Chitkara University, Baddi, Himachal Pradesh, India

² Department of Computer Science and Engineering, Malla Reddy Engineering College(A), Hyderabad, India

³ Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur (D.T), India

⁴ School of Computer Science and Artificial Intelligence, SR University, Warangal, Telangana 506371, India

⁵ Department of CSE, GMR Institute of Technology, Rajam, India

⁶ University Institute of Engineering-Computer Science and Engineering, Chandigarh University, Mohali, Punjab, India

with the best detection techniques is an initial and critical step that leads to the improvement of a patient's life [8]. X-rays and CT scans are the most commonly used imaging modalities for lung cancer detection [36].

CT scans are preferred over X-rays because it provides more information than an X-ray [33]. The earlier detection of LC from CT images [35] reduced the mortality rate by 20% in comparison with single view radiography [8]. The finest imaging approach CT imaging [9] is consistent for LC diagnosis because it has the potential to reveal each suspected as well as unsuspected LC nodules [1]. Nevertheless, intensity variations in CT images and anatomical formation are mistaken by doctors and radiologists which could cause complexity in marking the cancerous cell [3]. In recent times, for assisting radiologists and doctors in detecting cancer exactly, Computer Aided Diagnosis (CAD) is one of the most promising tools [10, 11]. Currently, CT is the major investigated imaging technique for detecting lung nodules owing to the fact that it has lower processing time; superior resolution; and eliminated artifacts [5]. The CT images and the presence of cancerous nodules in the lung are recognized by medical specialists as well as qualified radiologists [1]. A CT scan is a 3-D image of the body captured by a specialized machine that takes multiple images of the same anatomical area from different angles. Additionally, a CT scan aids in assessing intra thoracic pathological conditions [11]. This displays the information of the lung very clearly [12]. Such CT scan offers detailed images of the patient's chest to provide superior detection of LC [3, 13].

The enormous development of Deep learning (DL) techniques [37] in usual image detection and recognition has also changed diverse medical imaging issues and modalities [14]. DL-based Computer-Aided Diagnosis (CAD) systems can automatically extract higher-level features from original images using various network structures, surpassing classical CAD systems. However, they have some drawbacks, such as lower sensitivity and higher time consumption. The deep convolutional neural network (DCNN) technique is a DL model that significantly outperforms conventional methods in addressing medical image challenges. DL [15] is an essential image-segmentation approach also it is used for separating the internal boundaries of the lung in an effectual and quick manner [16]. DL-based LC screening approaches are capable of decreasing mortality by detecting the disease in the initial stage [17]. It can help in decreasing false negatives (FN) by detecting subtle or earlier indicators of LC that humans may leave unidentified [18]. Imaging modalities like CT, and Positive emission Tomography (PET) are incorporated by employing DL algorithms. DL module effectively analyzes CT scans, chest radiography, and further imaging modalities owing to their capability for evaluating massive data [19]. The DL-based lung imaging approaches mostly comprise detecting pulmonary nodules, segmentation, and

classification of pulmonary nodules. Nevertheless, DL approaches and numerous novel approaches and applications have developed every year. Utilizing DL techniques, physicians may achieve insights from a patient's state by merging data from numerous imaging modalities and other clinical data sources [10].

The major intention is to introduce Pyramid-KNet to detect LC employing CT images. Firstly, the CT images are passed to the image enhancement process employing Histogram equalization. After image enhancement, lung lobe segmentation is conducted employing Psi-Net. Thereafter, the lung nodule identification is performed on the basis of a grid-based strategy. Following this, feature extraction is accomplished to extract features like WLD, Textron, MRELBP with DCT, statistical features and GLCM features. Lastly, LC is detected utilizing Pyramid-KNet, which is the integration of PyramidNet and DKN.

The foremost contributions of the investigation are as follows

1.1 Proposed Pyramid-KNet for LC detection

- The main objective of the paper is to introduce a novel Pyramid-KNet method for LC detection.
- Here, the image enhancement is performed by Histogram equalization and lobe segmentation is conducted by employing Psi-Net.
- The LC is detected by employing Pyramid-KNet and is introduced through the amalgamation of DKN and PyramidNet.

The structure of this research is outlined as follows: Part 2 details former techniques and their limitations, Part 3 presents the proposed methodology and the architecture of Pyramid-KNet, Part 4 discusses the outcomes of Pyramid-KNet, and Part 5 concludes the findings.

2 Motivation

LC is an abnormal growth of cancerous tissue and 20% of patients have an entire survival rate of five years due to LC. However, detecting LC at earlier stages is very essential for people's survival. This motivates us to introduce a new technique for detecting LC by addressing the challenges of previous approaches.

2.1 Literature Survey

Tiwari et al. [1] presented Target based Weighted Elman DL Neural Network (TWEDLNN) model for LC detection. Here, the Modified Clip limit-based Contrasts Limited Adaptive Histograms Equalization (MC-CLAHE) algorithm

is employed for lung image enhancement and the new Mask-3FCM algorithm is utilized to detect the lung nodule. The model was capable of detecting the lung nodule during examination and screening but was affected by computational complexity. Alzubaidi et al. [3] developed a Local feature extraction framework for detecting LC employing CT images. This model utilizes various classification techniques for extracting features from these images. Even though, it had the potential to detect the suspicious regions inside the CT scan image, it was failed to include more data sets. Wahab Sait [10] devised an LC detection model for detecting LC. This model extracted critical features from the CT images. Although, it offered inadequate computational resources for detecting LC from the complicated image, it neglected to categorize the LC type with the lower quality image. Braveen et al. [7] introduced a novel Ant lion-based Auto encoders (ALbAE) for detecting LC. This module obtained less runtime, however it did not detect the stages of LC.

Shakeel et al. [6] established an improved Deep Neural network + ensemble classifier for detecting LC. Even though this module was exactly evaluating and interpreting difficult patterns in the images, it was suffered due to overfitting issues. Sori et al. [8] developed a “denoising first” two-path convolutional neural network (DFD-Net) for detecting LC utilizing CT images. This technique was useful for detecting the difference among nodule dimensions. This method was useful in detecting differences among nodule dimensions and was easily adaptable. However, the transparency that could potentially hinder trust and acceptance amongst healthcare professionals and patients was not considered in the model. Jena et al. [20] presented a Deep Gaussian mixture model in a region-based convolutional neural network (DGMM-RBCNN) for detecting LC. Here, a region-growing segmentation algorithm is employed to separate the lung region from its surroundings. Even though this model had the potential to process an enormous amount of data faster, it failed to enhance the diverse modality of LC images. Nazir et al. [16] established Image fusion for LC detection. The image fusion method utilizes Laplacian Pyramid (LP) decomposition combined with Adaptive Sparse Representation (ASR). This module decreased the image noise and enhanced the quality of the image, but it neglected to apply it to more databases.

2.2 Challenges

The experienced limitations by traditional techniques for LC detection are elucidated as below,

The module developed in [1] recognized even smaller nodules in their earlier stages and allowed for timely intervention and probably increased survival rates, even

though it failed to handle a variety of imaging data types and dimensions for ensuring performance across diverse databases.

In [6], the developed module neglected to evaluate complex CT images and did not differentiate the nodules accurately. Also, this module required more time to recognize the cancer.

The DFD-Net in [8] neglected to handle differences in image quality and diverse types of noise usually found in medical imaging data. Also, it did not validate and fine-tune the network to ensure its efficiency in detecting LC. LC is one of the most frequent diseases amongst humans and it is the major reason for increasing mortality. The poor prognosis of LC is due to ineffective diagnostic tools for early detection and the necessity for effective treatment.

The image enhancement is performed by using histogram equalization, which enhances the appearance of the images and provides clear images. Also, Psi-Net is used for lung lobe segmentation, which accurately detects the nodules and reduces the computational time of the model. Moreover, feature extraction improves the efficiency and accuracy of the model. Further, the integration of PyramidNet and DKN fine-tunes the model and efficiently detects LC. It is designed to handle various imaging data types and dimensions efficiently, ensuring consistent performance across different databases. Hence, the Pyramid-KNet model provides better performance than the traditional methods.

3 Proposed Pyramid-KNet for LC detection using CT images

LC is continuously considered as an important cause of death in cancer affected patients. Amongst the several disease examination processes, LC wants more concentration. Pyramid-KNet is introduced for detecting LC employing CT images. At first, the CT images acquired from the specific dataset [21] are fed into the image enhancement process, which is performed using Histogram equalization [22]. After image enhancement, lung lobe segmentation is carried out employing Psi-Net [23]. Thereafter, the lung nodule identification is performed employing the grid based strategy [24]. Following this, feature extraction is conducted for extracting features like WLD [25], Texton, and MRELBP [26] with DCT and statistical features including mean, variance, standard deviation, skewness, kurtosis and GLCM features [27]. At Last, LC detection is executed utilizing Pyramid-KNet, which is the combination of PyramidNet [28] and DKN [29] where layers are modified using Taylor concepts. The architectural diagram for Pyramid-KNet for LC detection is elucidated in Fig. 1.

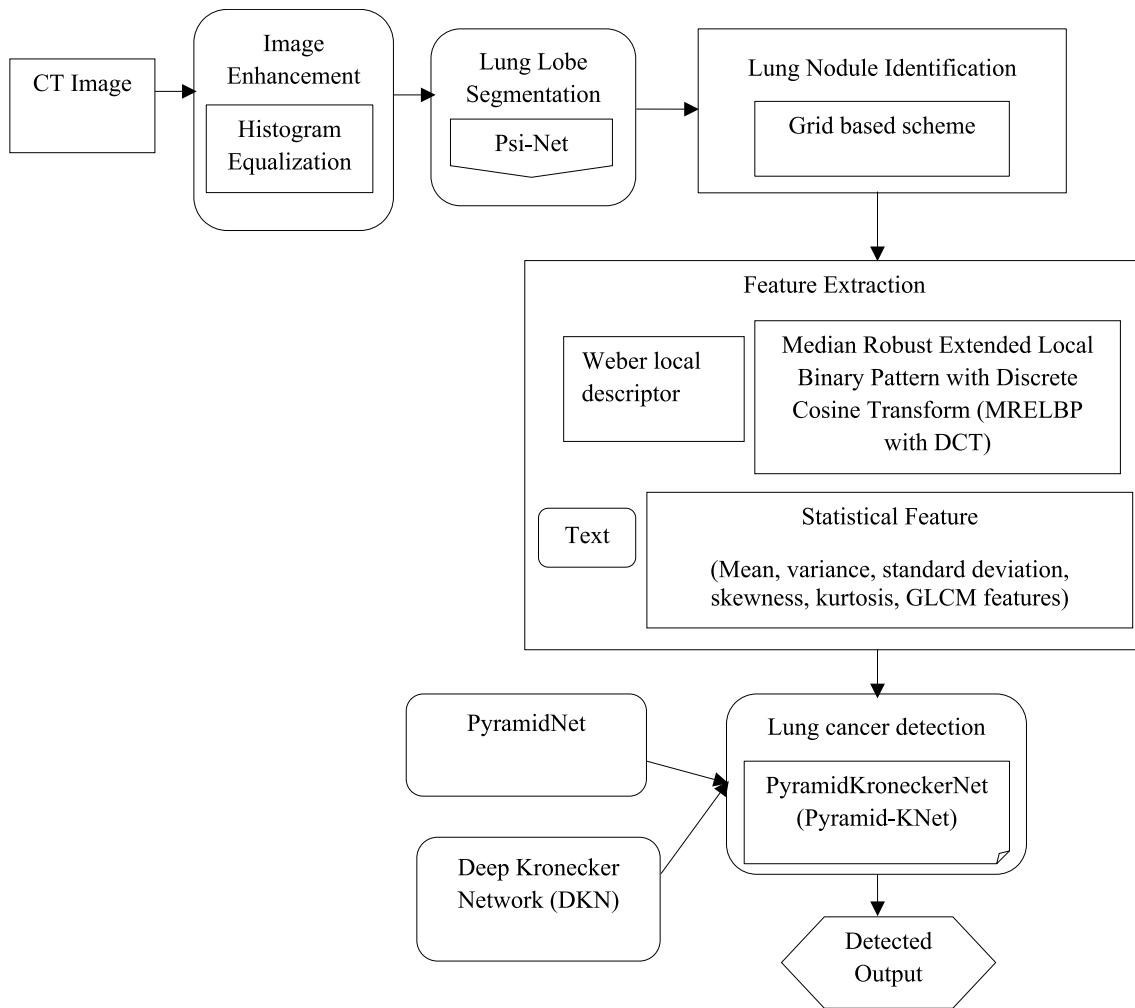


Fig. 1 The architecture of Pyramid-KNet for LC detection

3.1 Image acquisition

Image acquisition is a primary procedure for LC detection. Let us assume a database F for detecting LC with j a number of images and it is listed beneath,

$$F = \{F_1, F_2, \dots, F_o, \dots, F_j\} \quad (1)$$

Here, the input database is denoted as F whereas, the o^{th} image present in the database from the entire j samples is signified as F_o .

3.2 Image enhancement using histogram equalization

The input given to the image enhancement phase is F_o . It is a procedure for enhancing the image visual details and also aids in the detection of background information, which is

necessary for understanding the human viewers and offering the best input for further automated image processing approaches. Here, enhancement in the image is performed by histogram equalization.

Histogram equalization [22] refers to a familiar approach for improving image quality. Assume that the obtained image is mostly dark. Then its histogram must be skewed regarding the lowest end of the grey scale and the entire detail of the image is compressed into the histogram's dark end. Afterwards, at the dark end, stretch out the grey levels to generate a new uniformly dispersed histogram then the image must be clear. The resultant obtained from this phase is elucidated as M_o .

3.3 Lung lobe segmentation employing Psi-Net

The input M_o is passed to this phase. Lung lobe segmentation involves the identification and separation of individual lung lobes within a CT scan of the chest. This aids medical

professionals to examine and diagnose several lung-related diseases correctly. Here, lung lobe segmentation is carried out using Psi-Net.

3.3.1 Structure of Psi-Net

Psi-Net [23] refers to a UNet-based encoder-decoder network, with the same decoder paths on the right and a contracting encoder path on the left. It has the ability to recognize and separate lung lobes in medical images and this might lead to more exact segmentation and best diagnosis. The encoder path is composed of a repetitive down-sampling function.

Every down-sampling operation is prefaced through a convolution (CONV) function with 3×3 dimensions along with stride 1 that is succeeded through a ReLU. Every decoder block is symmetric for the encoder which aids in maintaining multiple features. The last CONV layer in the encoder is up-sampled and used as input for the decoder blocks.

Each decoder is tailored to perform various tasks. The blocks remain consistent in structure until the final layer, applied via 3×3 CONV, results in a single channel, which matches the output of input classes in the subsequent two blocks.

3.3.1.1 Loss function It incorporates three mechanisms: Contour decoder blocks, MSE loss for the distance decoder block, and Negative Log Likelihood (NLL) loss for the mask. Mask prediction is governed by predicting the distance map and contour, with the total loss calculated as follows:

$$E_{total} = \beta_1 E_{mask} + \beta_2 E_{contour} + \beta_3 E_{distance} \quad (2)$$

Here, $\beta_1, \beta_2, \beta_3$ determines scaling factors.

The individual losses are explained beneath,

(i) Mask: It distinguishes exact regions of interest inside an image, predominantly in medical imaging.

$$E_{mask} = \sum_{Y=\Omega} \log M_{mask}(Y; \hat{p}_{mask}(Y)) \quad (3)$$

Here, E_{mask} specifies categorization error in pixel-wise, Y demonstrates the position of the pixel in image space Ω . $M_{mask}(Y; \hat{p}_{mask})$ implies the predicted possibility for a true label E_{mask} following the softmax activation function.

(ii) Contour: It is a procedure of outlining or delineating the boundaries of a lung nodule in medical imaging, usually employing CT scans.

$$E_{contour} = \sum_{Y=\Omega} \log M_{contour}(Y; \hat{p}_{contour}(Y)) \quad (4)$$

Here, $E_{contour}$ specifies pixel-wise categorization error, whereas $\hat{p}_{contour}$ determines the true label.

(iii) Distance: It denotes the differentiation between the predicted outcome and the true target for quantifying individual prediction's quality.

$$E_{distance} = \sum_{Y=\Omega} (\hat{N}(Y) - N(Y))^2 \quad (5)$$

$E_{distance}$ indicates the pixel-wise MSE, $N(Y)$ implies the estimated distance map. X_o is the result obtained from this phase. The Psi-Net architecture is displayed in Fig. 2.

3.4 Lung nodule identification using grid based strategy

The input X_o is sent to the lung nodule identification phase to precisely locate the nodule regions. Lung nodule recognition is performed by adjusting the grid-based strategy [24]. This scheme efficiently identified the area of the affected nodule. A grid-based strategy divides the segments into blocks, called grids, to simplify calculations and reduce computation time. This approach segments the image into multiple grids to identify lung nodules, determining the area of the nodules. Therefore, the area of nodule identified is determined as N_o .

3.5 Feature extraction

Essential features for further processing are then extracted without losing important information. N_o is passed in this stage, Texton, WLD, and MRELBP with DCT features are extracted.

3.5.1 Texton

Assume there [30] are 2×2 and the values of pixels are denoted as Z_1, Z_2, Z_3, Z_4 . Even if the four pixel values are similar, then it is called texton. Consider, $K = (\mu, \nu)$ at the same location. Every image unit of texton comprises a pixel value and these image units of texton include five values of pixels, which is termed as P_1, P_2, P_3, P_4, P_5 . The result achieved from the texton feature is m_1 .

3.5.2 WLD

By employing the two concepts WLD [25] is created. The first concept determines differential excitation ψ and another one describes orientation Θ .

(i) Differential Excitation: The micro-difference within an image is calculated by the intensity differentiation between neighboring pixels, as detailed beneath,

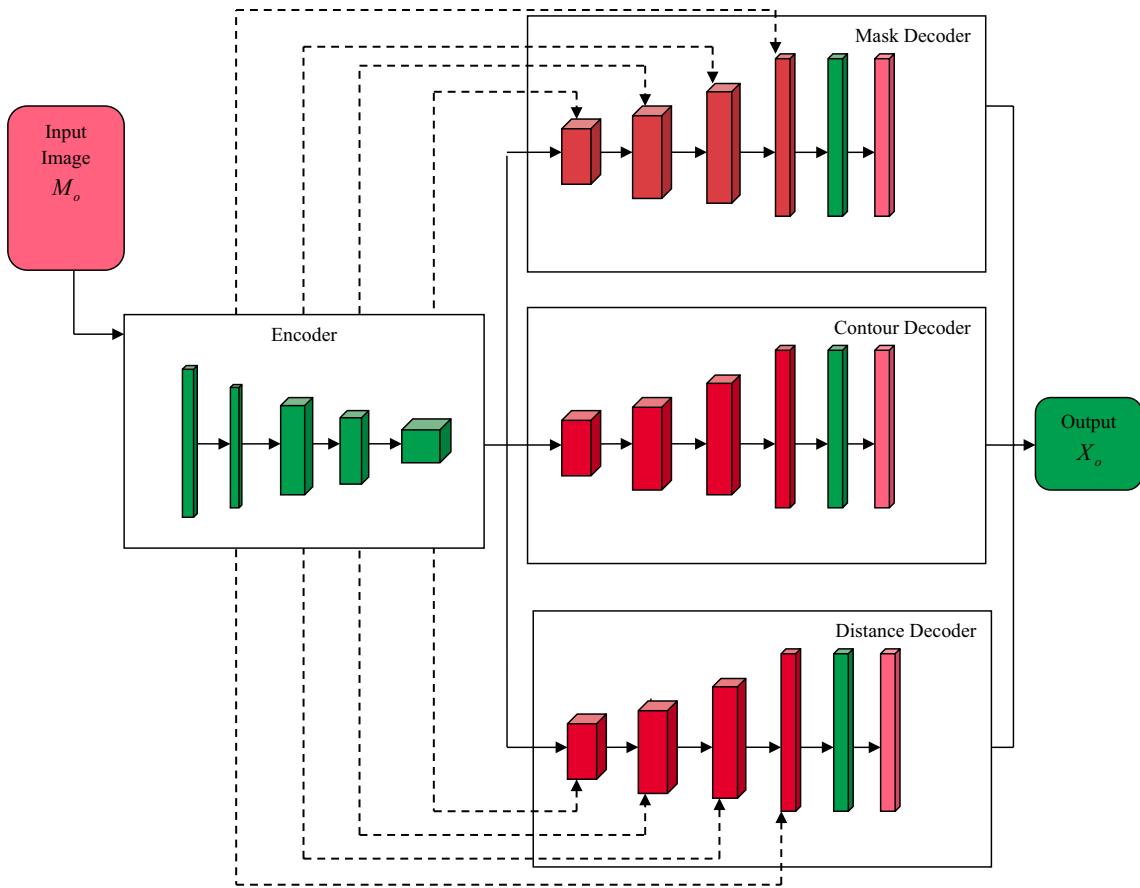


Fig. 2 Psi-Net architecture

$$Ev = \sum_{f=0}^{g=1} Ev(\sigma_x) = \sum_{f=0}^{g=1} v(\sigma_x) - v(\sigma_w) \quad (6)$$

Here, x^{th} neighbors of σ_w is determined through $\sigma_x (x = 0, 1, \dots, g - 1)$ if denotes the entire count of neighbors present in an area. Moreover, $v(\sigma_x)$ determines the neighbored pixel's intensity, and $v(\sigma_w)$ denotes the present pixel's intensity and it can be explained as follows,

$$\psi(\sigma_w) = \arctan \left(\frac{Ev}{v} \right) = \arctan \sum_{f=0}^{g=1} \left(\frac{v(\sigma_x) - v(\sigma_w)}{v(\sigma_x)} \right) \quad (7)$$

If $\psi(\sigma_w)$ is negative, afterwards, the present pixel is lighter by means of the neighbor pixels and also if $\psi(\sigma_w)$ is positive, afterwards the center pixel is darker by means of the neighbor pixels.

(ii) Orientation: It is defined as the pixel's directional property and it is calculated as follows,

$$\Theta(\sigma_w) = \arctan \left(\frac{Iv_\lambda}{Iv_N} \right) \quad (8)$$

Here, $Iv_\lambda = v(\sigma_7) - v(\sigma_3)$ and $Iv_N = v(\sigma_5) - v(\sigma_1)$ is computed from the two filters. The result obtained from the WLD feature is m_2 . The texton along with the WLD feature is denoted as,

$$m = \{m_1, m_2\} \quad (9)$$

Here, m_1 represents the texton feature, whereas m_2 specifies the WLD feature. Then, the statistical features and GLCM features are applied over m for obtaining the feature vector.

3.5.2.1 Statistical features The statistical features [31] like mean, variance, standard deviation, kurtosis and skewness are applied to the resultant m for producing a feature vector.

i) Mean: It [31] is computed by adding the pixel values from the image, which is divided by the entire pixels and it is expressed in the beneath expression,

$$\ell_1 = \sum_{N_o=0}^{X-1} N_o * P(N_o) \quad (10)$$

where, each gray level from the image m is signified as N_o , the probability of N_o is represented as $P(N_o)$, and the entire count of grey level is denoted as \mathbb{X} .

(ii) **Variance:** It [31] is defined as the dispersion of intensity values on the basis of the mean's value and it can be computed as follows,

$$\ell_2^2 = \sum_{N_o=0}^{\mathbb{X}-1} (N_o - \ell_1)^2 * O(N_o) \quad (11)$$

Here, ℓ_2 denotes variance while, ℓ_1 implies mean.

(iii) **Standard deviation:** It [31] is defined as the second central moment that displays the evaluation of homogeneity and it is computed as,

$$\ell_3 = \sqrt{\sum_{N_o=0}^{\mathbb{X}-1} (N_o - \ell_1)^2 * O(N_o)} \quad (12)$$

(iv) **Kurtosis:** It [31] calculates the leveling of distribution in regards to ordinary distribution and the expression is explained as beneath,

$$\ell_4^2 = \ell_1^{-4} \sum_{N_o=0}^{\mathbb{X}-1} (N_o - \ell_1)^2 * O(N_o) \quad (13)$$

(v) **Skewness:** The measure [31] of asymmetry over distribution by means of particular features around the mean. The skewness is computed as follows,

$$\ell_5^2 = \ell_1^{-3} \sum_{N_o=0}^{\mathbb{X}-1} (N_o - \ell_1)^3 * O(N_o) \quad (14)$$

The statistical features are denoted as SF_o , such that,

$$SF_o = \{\ell_1, \ell_2, \ell_3, \ell_4, \ell_5\} \quad (15)$$

3.5.2.2 GLCM features The GLCM [27] functions distinguish the image texture through computing how a frequent pair of pixels with particular values and in a particular spatial connection takes place in an image. By applying the GLCM feature over m for obtaining GL_o .

(i) **Contrast:** It [27] is defined as a measure of intensity contrast among a pixel and its neighbor to an entire image and is computed as below,

$$GL_1 = \sum_{ad,kl} |ad - kl|^2 / \rho(ad, kl)^2 \quad (16)$$

Here, the texture image is represented as ρ , whereas image grey level values are specified as ad and kl .

(ii) **Correlation:** It [27] is the measure of how correlated a pixel is to its neighbor to an entire image and it is computed as below,

$$GL_2 = \frac{(ad - \delta ad) / \rho(ad, kl)}{\xi_{ad} \xi_{kl}} \quad (17)$$

(iii) **Energy:** It [27] is defined as the sum of squared elements in the normalized GLCM and it can be expressed as,

$$GL_3 = \sum_{ad,kl} \rho(ad - kl)^2 \quad (18)$$

(iv) **Homogeneity:** It [27] refers to the value, which computes the nearness of the elements distribution in the GLCM over the GLCM diagonal.

$$GL_3 = \sum_{ad,kl} \frac{\rho(ad, kl)^2}{1 + ad - kl} \quad (19)$$

The GLCM feature is denoted as GL_o , such that,

$$GL_o = \{GL_1, GL_2, GL_3, GL_4\} \quad (20)$$

Then the SF_o and GL_o are applied over m to generate a feature vector C_1 .

$$C_1 = \{SF_o, GL_o\} \quad (21)$$

Here, SF_o specifies statistical features, whereas GL_o denotes GLCM features.

3.5.3 MRELBP feature with DCT feature

The input given to this phase is N_o . This section explains the MRELBP feature with DCT and the mathematical computations of this feature are given as follows:

3.5.3.1 MRELBP feature The MRELBP [26] has been broadly tested with diverse kinds of noise, like pixel arbitrary distortion Gaussian blur, and Gaussian white noise, and carries very proficient outcomes. Also, MRELBP has numerous significant features, along with fine features such as gray scale as well as rotation invariance. At last, MRELBP offers effectual outcomes quite simpler devoid of pre-training and adjusting parameters.

$$\alpha_{CI}(v_k) = \eta(\Phi(B_{k,\epsilon}) - \omega_\epsilon) \quad (22)$$

Here, α specifies the MRELBP feature, whereas $B_{k,\epsilon}$ denotes central and $\Phi(B_{k,\epsilon})$ illustrates the mean value for the area, which signifies the local patch v_k .

$$\alpha_{NI_{r,s}}(v_k) = \sum_{q=0}^{s-1} \eta(\Phi(B_{r,s,\epsilon_{r,q}}) - \omega_{r,s,\epsilon_r})^{2^q} \quad (23)$$

$$\omega_{r,s,\epsilon_r} = \frac{1}{s} \sum_{q=0}^{s-1} \eta \Phi(B_{r,s,\epsilon_{r,q}}) \quad (24)$$

Here, $B_{r,s}, \varepsilon_{r,q}$ demonstrates the local patch for the area, whereas r, s, ε_r, q denotes centered pixels for neighbors.

$$\alpha_RD_{r,s-1,\varepsilon_r,\varepsilon_{r-1}} = \sum_{q=0}^{s-1} \eta((\Phi(B_{r,s}, \varepsilon_r, q) - \Phi(B_{r-1,s}, \varepsilon_{r-1})))^{2^q} \quad (25)$$

The resultant obtained from the MRELBP feature is m_3 .

3.5.3.2 DCT feature DCT [32] is a kind of conversion that alters the image's dimension from the spatial field to the frequency field. DCT converts an image by summing cosine waves at various frequencies, reducing the image for quantization. In the frequency domain, DCT transforms the image into a gradient of sub-regions. The DCT transform is listed beneath,

$$N(\hbar, \ell) = \sigma(\hbar)\sigma(\ell) \sum_{ax=0}^{t-1} \sum_{by=0}^{t-1} v(ax, by) \cos \left[\frac{(2ax+1)\tau\pi}{2t} \right] \cos \left[\frac{(2by+1)\mu\pi}{2t} \right] \quad (26)$$

Here, $\tau, \mu = 0, 1, 2, \dots, t-1$, while t signifies the sequence dimensions, $t(ax, by)$ illustrates the spatial domain image as well as $N(\hbar, \ell)$ denotes the frequency domain of the image.

$$\sigma(\tau) = \begin{cases} \frac{1}{\sqrt{t}}, & \tau = 0 \\ \sqrt{\frac{2}{t}}, & \tau \neq 0 \end{cases} \quad (27)$$

Here, the DCT feature is applied to the MRELBP feature, which includes bands like Low-Low, High-Low, and High-High, and all these bands are X-ored to produce the result m_4 . The result obtained from MRELBP with DCT features is C_2 . Figure 3 illustrates the MRELBP with DCT features.

Finally, the obtained feature vector is expressed as below,

$$C_o = \{C_1 || C_2\} \quad (28)$$

3.6 Lung cancer detection utilizing Pyramid-KNet

LC is conducted through employing Pyramid-KNet, which is the incorporation of PyramidNet, as well as DKN. This

network comprises a DKN, Pyramid-KNet layer and PyramidNet approach. At first, input data F_o is given to DKN and then, the result attained from the DKN method A_o jointly with the feature vector C_o is passed to Pyramid-KNet which composes regression and fusion stages. Fusion is performed for integrating DKN and PyramidNet; the regression is conducted to confirm the similarity among predicted and targeted result. Lastly, the resultant from the Pyramid-KNet layer B_o along with F_o is subjected to the PyramidNet technique for additional process and the resultant D_o is achieved. Figure 4 shows the structure of the Pyramid-KNet model.

3.6.1 DKN model

DKN [29] is developed for assessing various medical imag-

ing data. DKN is capable of achieving the lowest sample size limitations, providing a preferred approach explanation and achieving prediction power as CNN. The input given to the DKN module is F_o .

Let us consider b samples by utilizing matrix represented images $K_e \in \beta^{r \times s}$ along with scalar responses R_e , $e = 1, \dots, u$. The responses R_e are assumed through the aid of a generalized linear technique, that is calculated as beneath,

$$A_o = \lambda(R_e) \exp \{R_e(F_o, I) - \rho(\langle F_o, I \rangle 1)\} \quad (29)$$

Here, $I \in \beta^{r \times s}$ signifies the target unrecognized coefficients matrix, $\lambda(\bullet)$ and $\rho(\bullet)$ is specified as an uni-variation function, whereas, the input image is denoted as F_o . This technique focuses on image estimation and ignores supplementary probable model variables, namely sex, age and so on.

They can be incorporated into the technique as needed. This approach has a specific known link function that is denoted as $\eta(\bullet)$,

$$\eta(\mathfrak{I}(R_e)) = \langle F_o, I \rangle \quad (30)$$

Utilizing the DKN structure aids in introducing coefficients I through rank- T Kronecker product decomposition with $Q(> 2)$ factors, as explained below.

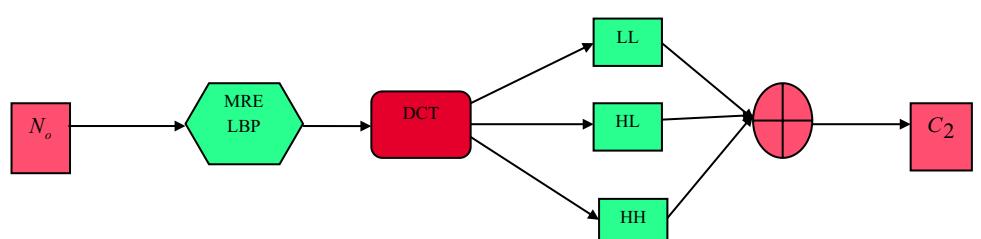


Fig. 3 MRELBP with DCT features

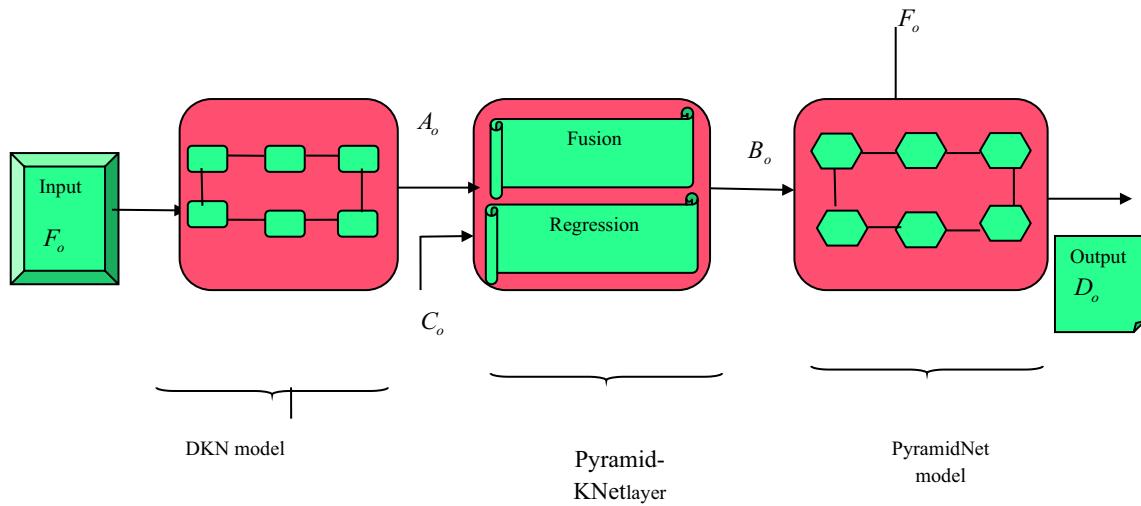


Fig. 4 Structure of Pyramid-KNet model

$$I = \sum_{n=1}^T V_Q^n \otimes V_{Q-1}^n \otimes \dots \otimes V_1^n \quad (31)$$

where, $V_q^n \in \beta^{r_q \times s_q}$, $q = 1, 2, 3, \dots, Q, n = 1, 2, \dots, T$ implies unknown matrices, while dimensions of V_w^n are not measured as known.

$$s = \prod_{w=1}^Q s_w \text{ as well as } \prod_{w=1}^Q y_w \quad (32)$$

The above expression can be re-written as,

$$V_{w'} = \otimes V_{w'-1} \otimes \dots \otimes V_{w''} = \otimes_{w=w'}^{w''} V_z \quad (33)$$

The outcome obtained through the DKN module is A_o . Figure 5 determines the Architecture of DKN.

3.6.2 Pyramid-KNet layer

Pyramid-KNet comprises phases such as fusion along with regression and the input subjected to this layer is G_o such that,

$$G_o = \{A_o, C_o\} \quad (34)$$

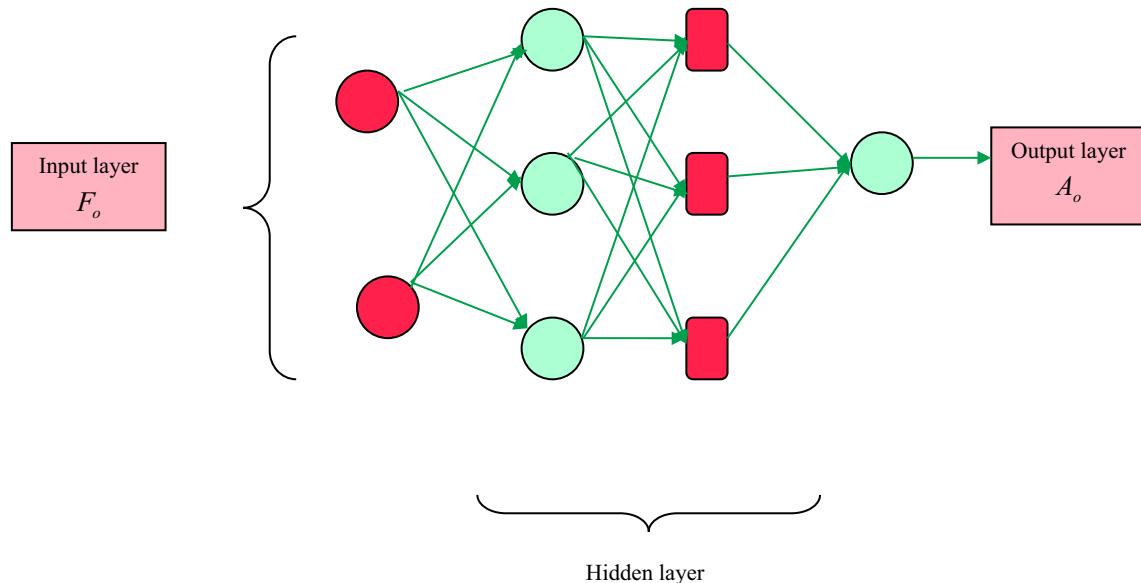


Fig. 5 Architecture of DKN

Here, the resultant obtained from the previous layer is denoted as A_o and C_o specifies the feature vector. The computation of Pyramid-KNet is listed below,

Applying Taylor concept,

$$J(d+1) = J(d) + \frac{J'(d)}{1!} \quad (35)$$

$$J'(d) = \frac{J(d) - J(d-q)}{q} \quad (36)$$

Assume, $q = 1$ and substitute $J'(d)$ in Eq. (34)

$$J(d+1) = J(d) + \frac{J(d) - J(d-1)}{1!} \quad (37)$$

$$J(d+1) = 2.J(d) - J(d-1) \quad (38)$$

Here, $J(d)$ specifies the resultant at d^{th} the interval, while $J(d-1)$ denotes the resultant at $(d-1)^{th}$ the interval.

$$B_o = 2 \cdot \sum_{p=1}^k C_{o(p)} * O_p - A_o \quad (39)$$

$$B_o = 2 \cdot \sum_{p=1}^k C_{o(p)} * O_p - \lambda(R_e) \exp \{R_e \langle K_e, I \rangle - \rho(\langle K_e, I \rangle)\} \quad (40)$$

Here, the entire count of the feature vector is indicated as k and O signifies the weight. The resultant achieved from this layer is denoted as B_o .

3.6.3 PyramidNet model

The main difference between PyramidNets [29] and further network structures is that the size of channels steadily enhances, because of persisting the size unless a residual unit by means of down sampling happens. The input subjected to this module is \mathcal{N}_o , such that,

$$\mathcal{N}_o = \{B_o, F_o\} \quad (41)$$

Here, B_o denotes the resultant from the preceding layer and F_o denotes the input image.

3.6.3.1 Feature map dimension configuration Initially, the calculation of feature map sizes E_h of the h^{th} residual units that refer to the m^{th} group is explained as follows,

$$E_h = \begin{cases} 16, & \text{if } m(h) = 1 \\ 16 \cdot 2^{m(h)-2}, & \text{if } m(h) \geq 2 \end{cases} \quad (42)$$

where, the group index for the h^{th} residual unit belongs is denoted as $m(h) \in \{1, 2, 3, 4\}$. The residual unit that belongs

to a similar group possess an equivalent feature map dimension, then the m^{th} group comprises H_m residual units. At the first group, it contains a CONV layer that transforms an RGB image into several feature maps. After that, H_m residual units are subjected, then the size of the feature is down sampled through half and the count of size is doubled for m^{th} group. Then, a technique for enhancing the feature map size is expressed as beneath,

$$E_h = \begin{cases} 16, & \text{if } h = 1 \\ \lfloor E_{h-1} + \chi/H \rfloor, & \text{if } 2 \leq h \leq H+1, \end{cases} \quad (43)$$

Here, H specifies the entire count of residual units, which is denoted as, $H = \sum_{m=2}^4 H_m$ whereas, the size is enhanced through a step factor χ/H . Moreover, the resultant dimensions of every group last unit turn to be $16 + (m-1)\chi/3$ with similar count of residual units.

The above expressions rely on the addition-based widening step factor χ for increasing sizes. Subsequently, Eq. (42) can be converted as below:

$$E_h = \begin{cases} 16, & \text{if } h = 1 \\ \left\lfloor E_{h-1} \cdot \chi^{\frac{1}{H}} \right\rfloor, & \text{if } 2 \leq h \leq H+1, \end{cases} \quad (44)$$

The key difference between additive and multiplicative PyramidNets is that the feature map size of an additive network grows linearly, whereas in a multiplicative network, it grows geometrically. Specifically, the size increases gradually in input-side layers and significantly in output-side layers.

3.6.3.2 Building block To enhance the network structure's capabilities, it is crucial to model a superior building block. PyramidNet aims to address this differently by gradually increasing the feature map size instead of doubling it at specific residual units, thus evenly distributing the task of enhancing the feature maps. PyramidNet effectively addresses this issue while simultaneously improving overall performance.

3.6.3.3 Zero-padded shortcut connection The PyramidNet did not utilize any shortcut, because it has the potential to identify the mapping alone, because the size of feature maps diverges amongst particular residual units. Hence, only a projection or zero-padded is employed for the entire residual units.

The result of the zero-padded identity mapping shortcut on h^{th} the residual unit, which $\in m^{th}$ group by means of the reshaped vector \mathfrak{R}_h^γ of the γ^{th} feature map:

$$\mathfrak{R}_h^\gamma = \begin{cases} \mathfrak{N}_{(h,\gamma)}(\mathfrak{R}_{h-1}^\gamma) + \mathfrak{R}_{h-1}^\gamma, & \text{if } 1 \leq \gamma \leq E_{h-1} \\ \mathfrak{N}_{(h,\gamma)}(\mathfrak{R}_{h-1}^\gamma) & \text{if } E_{h-1} < \gamma \leq E_h \end{cases} \quad (45)$$

Here, $\mathfrak{R}_{h-1}^\gamma$ defines the previous layer outcome B_o , whereas the γ^{th} residual operation of h^{th} the residual unit is denoted as $\mathfrak{N}_{(h,\gamma)}(\cdot)$. Furthermore, E_h determines the size of predefined channels of h^{th} residual units. From Eq. (43) the zero padded constituents are employed for enhancing the sizes. Let \mathfrak{R}_h^γ comprises the resultants of both plain networks and residual networks.

3.6.3.4 New building block In this phase, Batch Normalization (BN) layers, and the Rectified Linear Units (ReLUs) are considered for further processing and it is explained as beneath.

(i) ReLUs in a Building Block: Nonlinearity is crucial; however, using ReLUs after adding residual units adversely affects performance.

$$\mathfrak{R}_h^\gamma = \text{RELU}(\mathfrak{N}_{(h,\gamma)})(\mathfrak{R}_{h-1}^\gamma) + (\mathfrak{R}_{h-1}^\gamma) \quad (46)$$

Here, the ReLUs act like filters for non-negative elements, and pre-activation ResNets address this by placing pre-activated residual units with BN layers and ReLUs before the CONV layers.

$$D_o = \mathfrak{N}_{(h,\gamma)}(\mathfrak{R}_{h-1}^\gamma) + (\mathfrak{R}_{h-1}^\gamma) \quad (47)$$

$$D_o = (\mathfrak{N}_{h,\gamma})(2 \cdot \sum_{p=1}^k C_{o(p)} * O_p - \lambda(R_e) \exp \{R_e \langle K_e, I \rangle - \rho(\langle K_e, I \rangle)\}) + 2 \cdot \sum_{p=1}^k C_{o(p)} * O_p - \lambda(R_e) \exp \{R_e \langle K_e, I \rangle - \rho(\langle K_e, I \rangle)\} \quad (48)$$

Here, $\mathfrak{R}_{h-1}^\gamma$ defines the previous layer outcome B_o . The architecture of pyramidNet is displayed in Fig. 6.

4 Result and discussions

Pyramid-KNet is established for LC detection using CT images, aiming to achieve superior results. The evaluation confirms Pyramid-KNet's effectiveness, as detailed in this section.

4.1 Experimental setup

Pyramid-KNet, implemented in Python, is established for detecting LC with CT images. The parameter details of the developed approach and other existing methods are illustrated in Table 1.

4.2 Dataset description

The Lung Image Database (LIDC-IDRI) [21] is a web-accessible international resource for progression, training, and assessment of CAD techniques for detecting and diagnosing LC. Eight medical imaging companies created this database

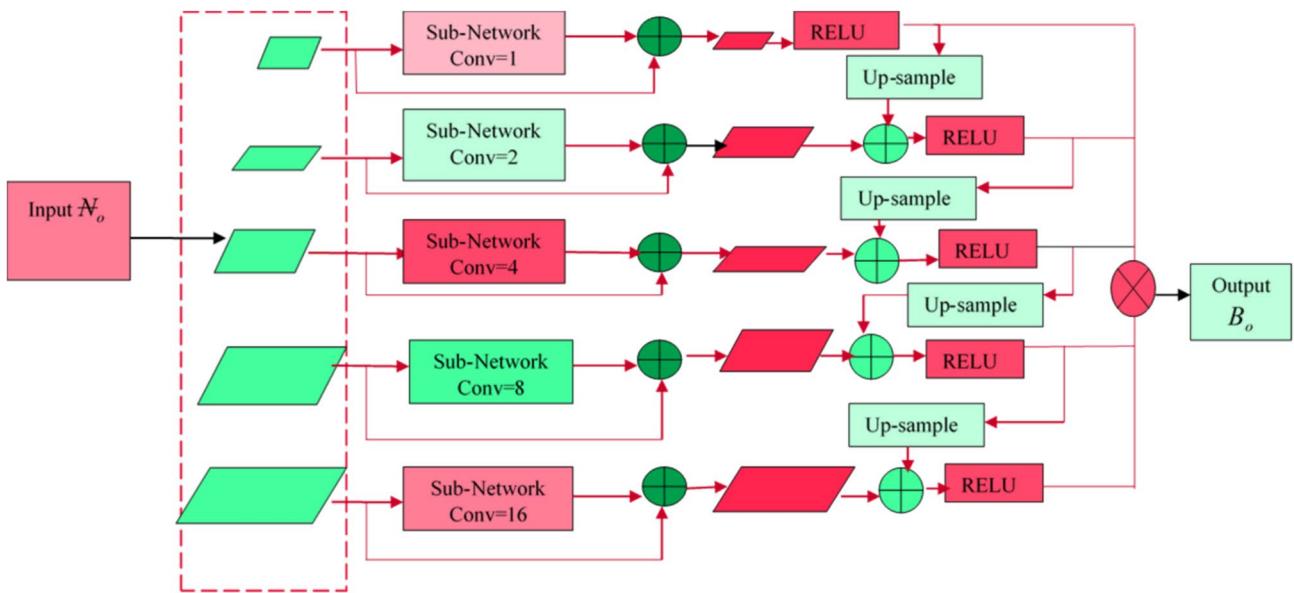
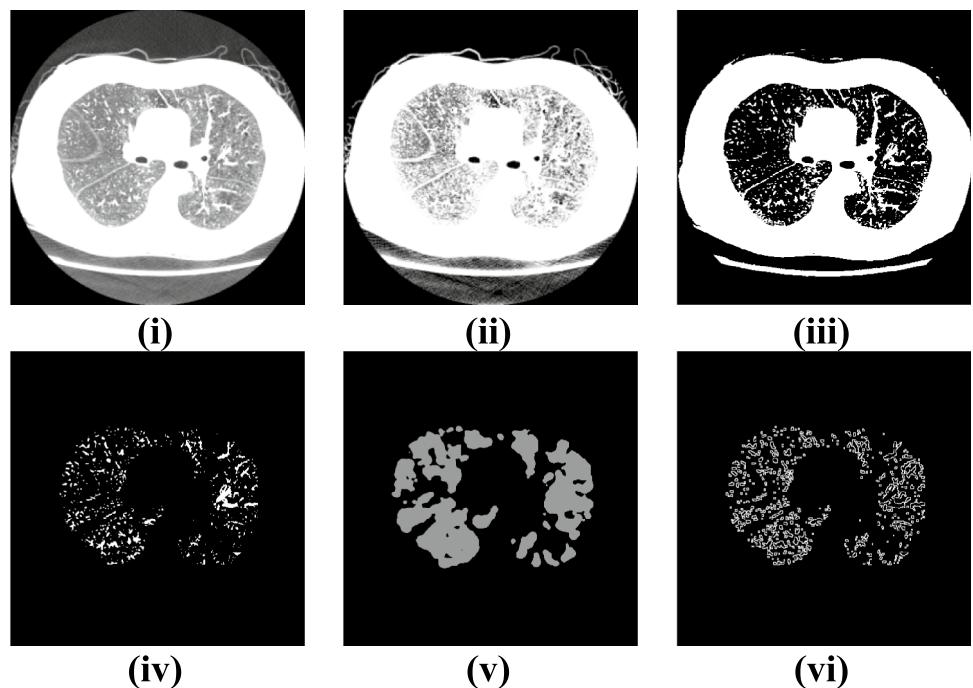


Fig. 6 Architecture of pyramidNet

Table 1 Experimental parameter

| Methods | Epoch | Batch size | Learning rate | Loss |
|---------------------------|-------|------------|---------------|--------------------------|
| 3D-DLCNN | 100 | 32 | 0.001 | categorical_crossentropy |
| CapsNet | 100 | 32 | 0.001 | categorical_crossentropy |
| Co-ReTr | 100 | 32 | 0.001 | categorical_crossentropy |
| DKN | 100 | 64 | 0.001 | categorical_crossentropy |
| TWEDLNN + 3FCM | 100 | 32 | 0.001 | categorical_crossentropy |
| ALbAE | 100 | 16 | 0.001 | categorical_crossentropy |
| DNN + Ensemble classifier | 100 | 32 | 0.001 | categorical_crossentropy |
| DFD-Net | 100 | 64 | 0.001 | categorical_crossentropy |
| PyramidNet | 100 | 32 | 0.001 | categorical_crossentropy |
| Proposed Pyramid-KNet | 100 | 32 | 0.001 | categorical_crossentropy |

Fig. 7 Image outcomes of Pyramid-KNet **i** input image, **ii** enhanced image, **iii** segmented image, **iv** lung nodule image, **v** texton image, **vi** WLD image

which comprises 1018 cases. A total of 700 CT cases are used in the experiment.

4.3 Experimental results

Figure 7 illustrates the image outcomes of Pyramid-KNet. Figure 7*i* depicts the input image, Fig. 7*ii* shows the enhanced image, Fig. 7*iii* presents the segmented image, Fig. 7*iv* displays the lung nodule image, Fig. 7*v* shows the texton image and Fig. 7*vi* details the WLD.

4.4 Estimation measures

The measures employed for Pyramid-KNet estimation are F-measure, accuracy, and precision.

4.4.1 Accuracy

It refers to the computation of accurate estimations, which is assessed in the expression below,

$$\text{Accuracy} = \frac{IJ + IM}{IJ + SS + IM + SK} \quad (49)$$

Here, the True positive is denoted as IJ , the True negative is indicated as IM , the false negative is represented as SS , and the false positive is specified as SK .

4.4.2 Precision

It is determined as the closeness of more than two dimensions along with one another and the expression is defined as follows,

$$KL = \frac{IJ}{IJ + SK} \quad (50)$$

4.4.3 F-measure

It refers to the harmonic mean of precision and recall, calculated by the following expression,

$$LR = \frac{2(KL * MJ)}{KL + MJ} \quad (51)$$

4.5 Performance evaluation

The performance estimation of Pyramid-KNet for detecting LC is based on varying the K-fold and training data through iterations.

4.5.1 Performance evaluation regarding training data

Figure 8 determines the estimation of Pyramid-KNet with numerous iterations through varying training data as 90% are computed. Figure 8i displays the estimation of Pyramid-KNet in consideration of accuracy. The accuracy obtained by Pyramid-KNet is 0.859 by iteration 20, 0.865 by iteration 40, 0.878 by epoch 60, 0.899 by iteration 80 and 0.922 by iteration 100. Figure 8ii deliberates the assessment of Pyramid-KNet considering precision. Pyramid-KNet gained precision of 0.841, 0.861, 0.873, 0.885 and 0.919 from 20 to 100 iterations. Figure 8iii determines computing Pyramid-KNet regarding F-measure. The values gained through Pyramid-KNet from 20 to 100 iterations are 0.865, 0.879, 0.885, 0.908 and 0.936.

4.5.2 Performance assessment regarding K-fold

Figure 9 explains the assessment of Pyramid-KNet through different iterations by altering the K-fold as 9. Figure 9i

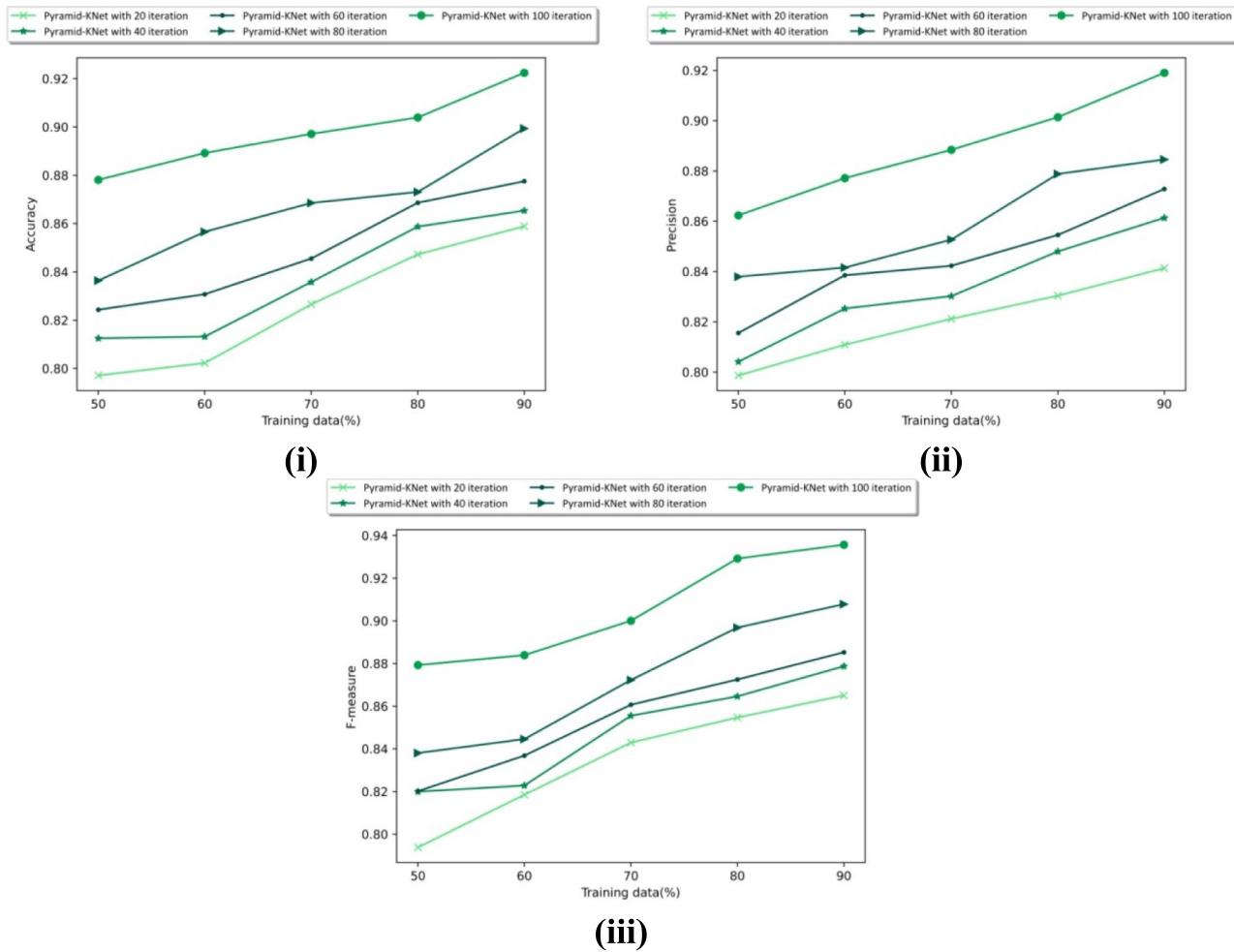


Fig. 8 Performance evaluation regarding training data, **i** accuracy, **ii** precision, **iii** F-measure

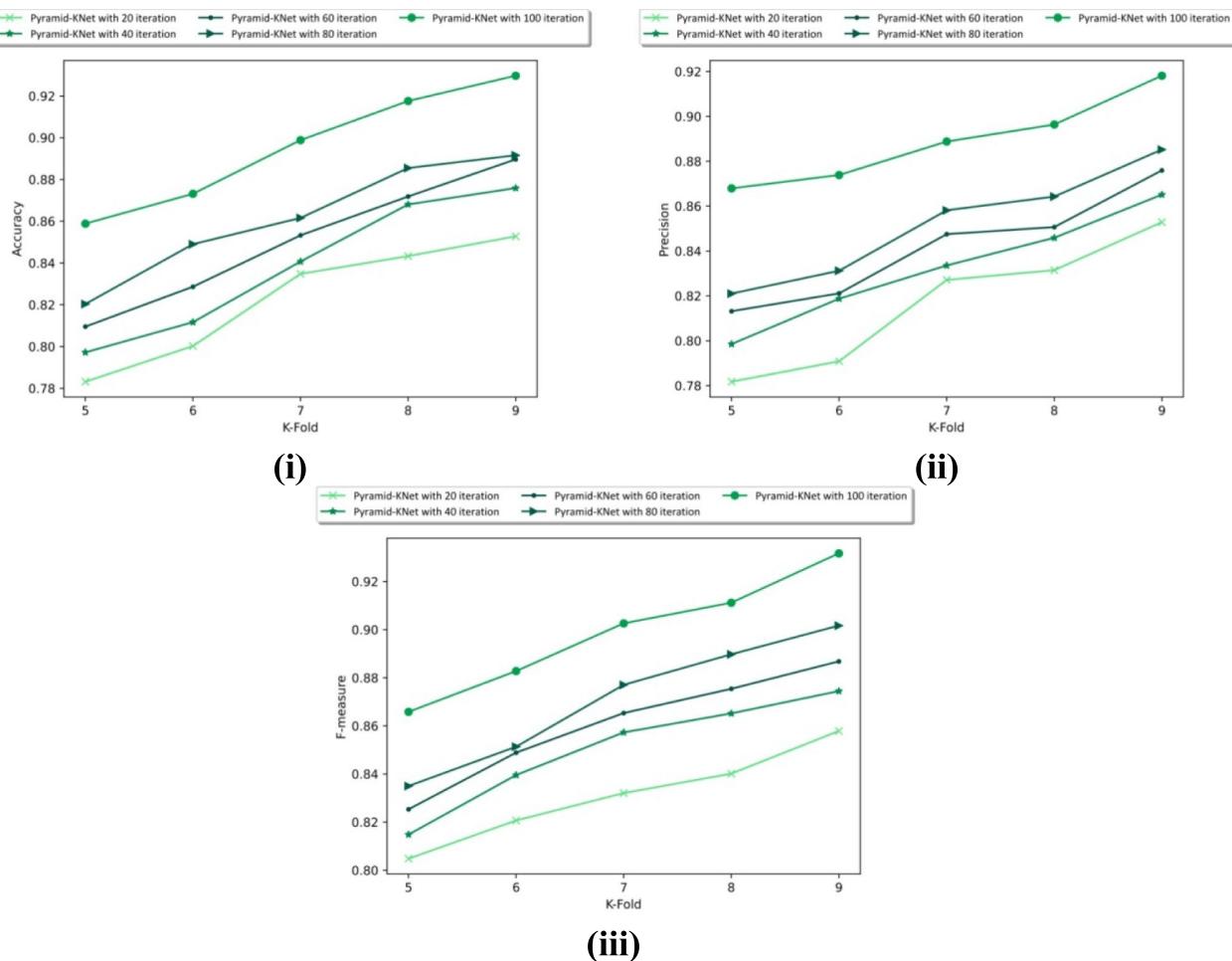


Fig. 9 Performance regarding K-fold, **i** accuracy, **ii** precision, **iii** F-measure

elucidates the evaluation of Pyramid-KNet in regards to accuracy. Accuracy acquired through Pyramid-KNet are 0.853, 0.876, 0.890, 0.892 and 0.930 from 20 to 100 iterations. Figure 9ii elucidates the computation of Pyramid-KNet regarding precision. The values attained by Pyramid-KNet from 20 to 100 iterations are 0.853, 0.865, 0.876, 0.885 and 0.918. Figure 9iii deliberates the estimation of Pyramid-KNet concerning F-measure. Pyramid-KNet obtained F-measure of 0.858, 0.874, 0.887, 0.902 and 0.932 from 20 to 100 iterations.

4.6 Comparative approaches

TWEDLNN + 3FCM [1], ALbAE [7], DNN + Ensemble classifier [6], DFD-Net [8], Three-Dimensional Deep Learning Convolutional Neural Network (3D-DLCNN) [40], Capsule Network (CapsNet) [38], Convolutional neural networks, Residual blocks, and Transformers (Co-ReTr) [39], DKN [29], and PyramidNet [28] are the approaches considered to assess with Pyramid-KNet for revealing its efficacy.

4.7 Comparative evaluation

The comparative evaluation of Pyramid-KNet for detecting LC is achieved by varying the training data and the K-fold.

4.7.1 Analysis of Pyramid-KNet with training data

The evaluation of Pyramid-KNet through varying training data for considered metrics is shown in Fig. 10. This section analyzes the values obtained by previous techniques and Pyramid-KNet, using 90% of the training data. The computation of Pyramid-KNet in regard to accuracy is elucidated in Fig. 10i. The accuracy obtained by Pyramid-KNet is 0.929, whereas the values gained by 3D-DLCNN are 0.815, CapsNet is 0.859, Co-ReTr is 0.846, DKN is 0.866, TWEDLNN + 3FCM is 0.886, ALbAE is 0.833, DNN + Ensemble classifier is 0.852, DFD-Net is 0.878, and PyramidNet is 0.906. The superior performance of Pyramid-KNet to that of existing approaches is 12.272%, 7.531%, 8.934%, 6.782%, 4.628%,

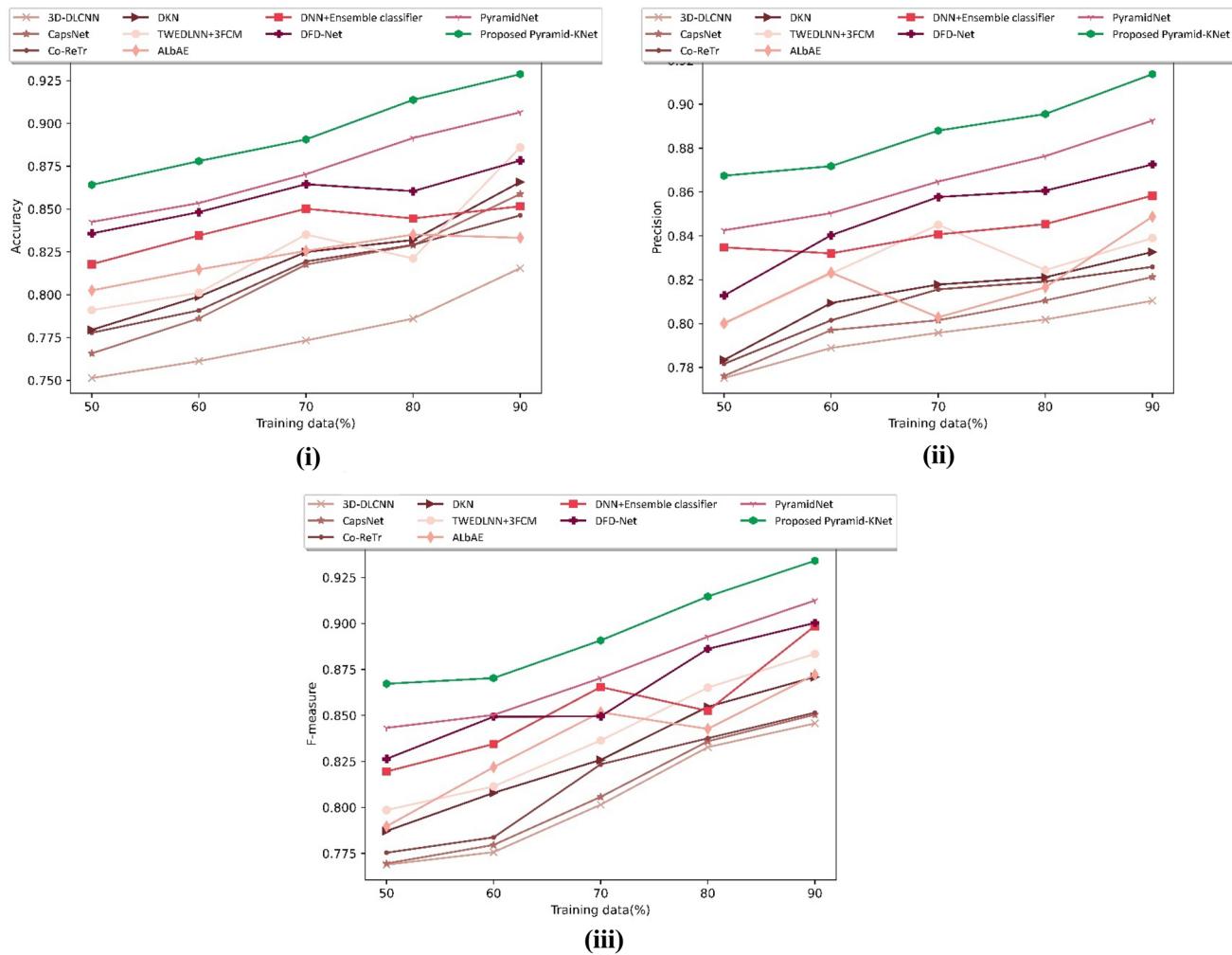


Fig. 10 Analysis of Pyramid-KNet with training data, **i** accuracy, **ii** precision, **iii** F-measure

10.289%, 8.307%, 5.434%, and 2.476%. Figure 10ii demonstrates an evaluation of Pyramid-KNet in regards to precision. Pyramid-KNet accomplished a precision of 0.914 whilst 3D-DLCNN, CapsNet, Co-ReTr, DKN, TWEDLNN + 3FCM, ALbAE, DNN + Ensemble classifier, DFD-Net, and PyramidNet, accomplished 0.810, 0.821, 0.826, 0.833, 0.839, 0.849, 0.858, 0.873, and 0.893. It attained performance gains of about 13.020%, 11.816%, 11.269%, 10.503%, 8.183%, 7.111%, 6.061%, 4.506%, and 3.939%. Figure 10iii shows the evaluation of Pyramid-KNet with respect to the F-measure. 3D-DLCNN, CapsNet, Co-ReTr, DKN, TWEDLNN + 3FCM, ALbAE, DNN + Ensemble classifier, DFD-Net, and PyramidNet, gained F-measure of 0.846, 0.850, 0.852, 0.871, 0.883, 0.872, 0.899, 0.900, and 0.912 while Pyramid-KNet attained 0.934. The enhancement in performance of proposed to that of former approaches is 9.422%, 8.994%, 8.779%, 6.745%, 5.415%, 6.629%, 3.797%, 3.615%, and 2.355%.

4.7.2 Analysis of Pyramid-KNet with K-fold

Figure 11 shows the estimation of Pyramid-KNet through varying K-fold for considered measures. This section discusses the values achieved by previous techniques and Pyramid-KNet when applying a ninefold cross-validation. Assessment of Pyramid-KNet in regards to accuracy is explained in Fig. 11i. 3D-DLCNN, CapsNet, Co-ReTr, DKN, TWEDLNN + 3FCM, ALbAE, DNN + Ensemble classifier, DFD-Net, and PyramidNet gained accuracy of 0.841, 0.844, 0.849, 0.859, 0.870, 0.880, 0.864, 0.892, and 0.906 whilst Pyramid-KNet attained 0.927. The enhancement in performance of proposed to that of traditional techniques are 9.277%, 8.954%, 8.414%, 7.335%, 6.144%, 5.053%, 6.841%, 3.794%, and 2.265%. Figure 11ii elucidates the computation of Pyramid-KNet concerning precision. Precision accomplished by Pyramid-KNet is 0.916, whereas the values gained by 3D-DLCNN

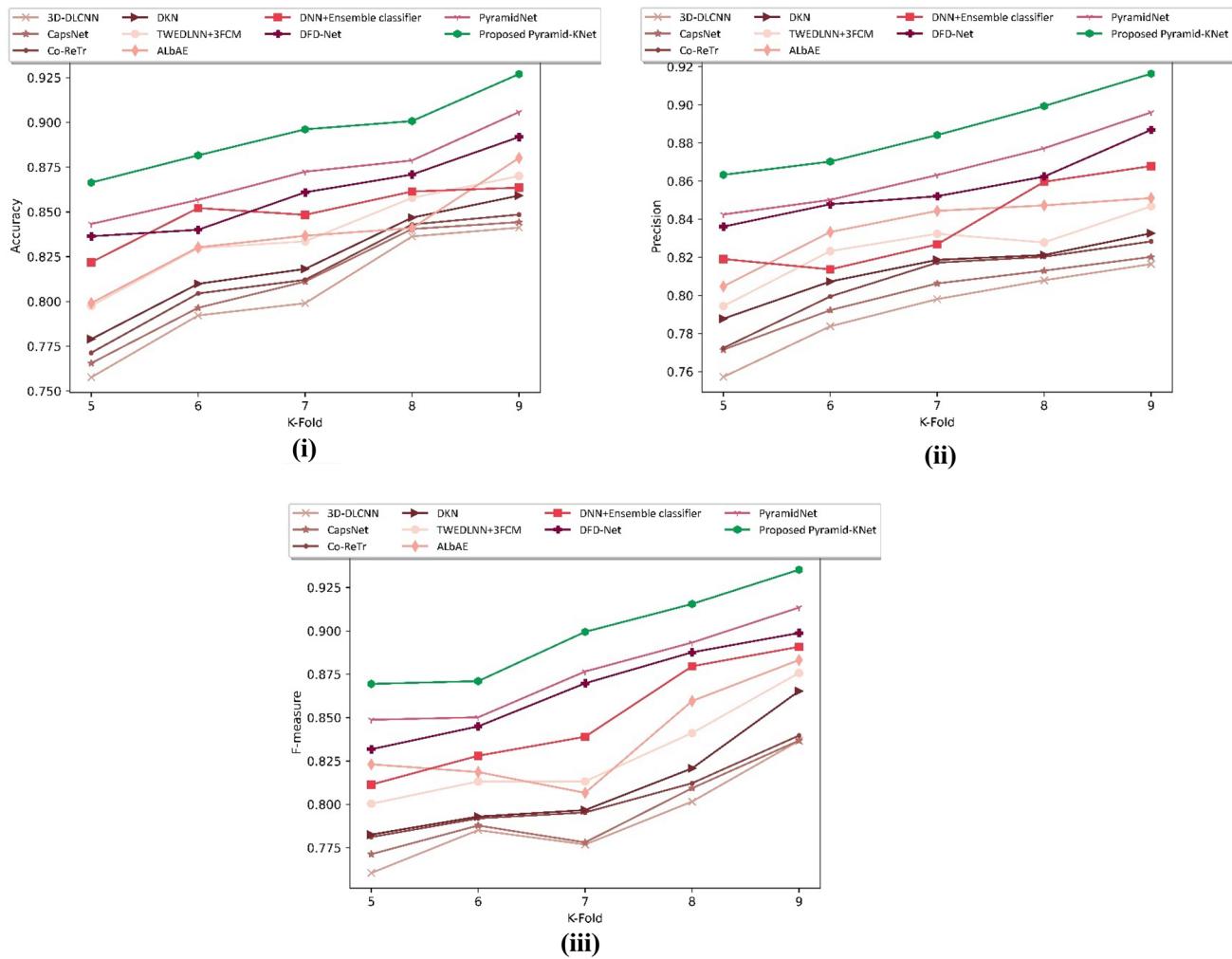


Fig. 11 Analysis of Pyramid-KNet with K-fold, **i** accuracy, **ii** precision, **iii** F-measure

is 0.816, CapsNet is 0.820, Co-ReTr is 0.828, DKN is 0.833, TWEDLNN + 3FCM are 0.847, AlbAE is 0.851, DNN + Ensemble classifier is 0.868, DFD-Net is 0.887, and PyramidNet is 0.896, The performance improvement of Pyramid-Knet to that of classical techniques is 10.917%, 10.480%, 9.607%, 9.061%, 7.596%, 7.123%, 5.291%, 3.220%, 2.183%. Figure 11iii determines the estimation of Pyramid-Knet in regards to F-measure. Pyramid-Knet accomplished an F-measure of 0.935, whilst 3D-DLCNN, CapsNet, Co-ReTr, DKN, TWEDLNN + 3FCM, AlbAE, DNN + Ensemble classifier DFD-Net, and PyramidNet, accomplished 0.837, 0.837, 0.840, 0.865, 0.876, 0.883, 0.891, 0.899, and 0.913. It achieved performance gains of about 10.481%, 10.481%, 10.160%, 7.487%, 6.363%, 5.564%, 4.750%, 3.897%, 2.353%.

4.8 Comparative discussion

Table 2 shows the comparative discussion. Pyramid-KNet achieved better results compared to the existing modules like 3D-DLCNN, CapsNet, Co-ReTr, DKN, TWEDLNN + 3FCM, AlbAE, DNN + Ensemble classifier, DFD-Net, and PyramidNet, by altering k-fold. Pyramid-KNet attained an accuracy of 93%, while the existing modules like 3D-DLCNN, CapsNet, Co-ReTr, DKN, TWEDLNN + 3FCM, AlbAE, DNN + Ensemble classifier and DFD-Net, and PyramidNet gained 84%, 84%, 85%, 86%, 87%, 88%, 86%, 89%, and 91%. From this, it implies that the Pyramid-KNet module had the potential to detect particular infections more precisely than the prior models. The precision value gained by Pyramid-KNet is

Table 2 Comparative discussion of Pyramid-KNet

| Analysis based upon | Metrics/methods | 3D-DLCNN | CapsNet | Co-ReTr | DKN | TWEDLNN+3FCM | AlbAE | DNN+Ensemble classifier | DFD-Net | PyramidNet | Proposed-Pyramid-KNet |
|----------------------------|-----------------|----------|---------|---------|-----|--------------|-------|-------------------------|---------|------------|-----------------------|
| Training Data = 90% | | | | | | | | | | | |
| K-fold = 9 | Accuracy | 82% | 86% | 85% | 87% | 89% | 83% | 85% | 88% | 91% | 93% |
| | Precision | 81% | 82% | 83% | 84% | 83% | 85% | 86% | 87% | 89% | 91% |
| | F-measure | 85% | 85% | 87% | 88% | 87% | 87% | 90% | 90% | 91% | 93% |
| K-fold = 9 | Accuracy | 84% | 84% | 85% | 86% | 87% | 88% | 86% | 89% | 91% | 93% |
| | Precision | 82% | 82% | 83% | 83% | 83% | 85% | 87% | 89% | 90% | 92% |
| | F-measure | 84% | 84% | 84% | 87% | 88% | 88% | 89% | 90% | 91% | 94% |

92%, whereas the values gained by traditional techniques 3D-DLCNN is 82%, CapsNet is 82%, Co-ReTr is 83%, DKN is 83%, TWEDLNN + 3FCM are 85%, AlbAE is 85%, DNN + Ensemble classifier is 87% and DFD-Net is 89%, and PyramidNet is 90%. It implies that the proposed module had the ability to identify the cancer affected image exactly. F-measures attained by existing modules are 84%, 84%, 84%, 87% 88%, 88%, 89%, 90%, and 91% whereas the Pyramid-KNet attained 94% of F-measure. The intended Pyramid-KNet obtained accuracy of 93%, 92% of precision and 94% of F-measure. Hence, the Pyramid-KNet has proven to be an effective strategy for LC detection.

The reasons for the better performance of the proposed Pyramid-KNet are discussed as follows:

In this research, image enhancement is performed using histogram equalization, which improves the visibility of the images. It provides accurate information from the enhanced images, which increases the accuracy of the model. Then, lung lobe segmentation is done by Psi-Net. Psi-Net maintains shape well with better boundary outputs and enhances segmentation performance. Feature extraction reduces the complexity of the data by focusing on the most informative aspects. PyramidNet provides accurate and reliable predictions. DKN has advantages like high convergence speed, minimal computational error and cost. Hence, the hybridization of DKN and PyramidNet performs well in lung cancer detection.

4.9 Limitation and future work

The major limitation of the model is that it requires high computational resources, which limits its accessibility, particularly in resource-constrained clinical environments. Future investigations will focus on comparison between X-ray and CT scan images for superior outcomes of detecting LC. Also, future work will focus on expanding the training datasets to include a wider variety of lung cancer types, imaging conditions, and demographic factors to enhance generalizability.

5 Conclusion

LC is a severe disease characterized by the uncontrolled growth of harmful cells in the lungs. Early detection significantly increases the chances of effective treatment before the disease reaches a critical stage. Recently, novel technologies have attained more concentration through, the tumor's appearance, identifying their presence, shape, size and higher discrepancy is a demanding task. Here, Pyramid-KNet is established for detecting LC employing CT images. Initially, CT images from specific databases undergo image enhancement using histogram equalization.

After that, by employing using Psi-Net lung lobe segmentation is performed. Then, the lung nodule identification is performed employing the grid-based strategy. Furthermore, feature extraction is carried out for extracting various features. Finally, LC detection is executed employing Pyramid-KNet, which is the amalgamation of PyramidNet and DKN, where layers are modified employing Taylor concepts. The intended Pyramid-KNet gained superior accuracy of 93%, 92% of precision and 94% of F-measure.

Author contributions Satish Muppudi and Praveen Kantha was responsible for the conceptualization, data collection, part of the methodology, prepared figures and wrote most part of the manuscript. Katakam Venkateswara Rao and J Anitha was responsible for the model training and evaluation, diagram designation, and writing part of the manuscript. Satish Muppudi, Balajee Maram, and Dr. Parul Datta were responsible for project supervision, result interpretation, ethical considerations and editing manuscript for final submission.

Data availability The dataset utilized throughout this research project is publicly available via the link: <https://wiki.cancerimagingarchive.net/display/Public/LIDC-IDRI>.

Declarations

Conflict of interest The authors declare no competing interests.

References

- Tiwari L, Raja R, Awasthi V, Miri R, Sinha GR, Alkinani MH, Polat K (2021) Detection of lung nodule and cancer using novel Mask-3 FCM and TWEDLNN algorithms. Measurement 172:108882
- National Lung Screening Trial Research Team (2011) Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med* 365(5):395–409
- Alzubaidi MA, Otoom M, Jaradat H (2021) Comprehensive and comparative global and local feature extraction framework for lung cancer detection using CT scan images. *IEEE Access* 9:158140–158154
- Suzuki K, Kusumoto M, Watanabe SI, Tsuchiya R, Asamura H (2006) Radiologic classification of small adenocarcinoma of the lung: radiologic-pathologic correlation and its prognostic impact. *Ann Thorac Surg* 81(2):413–419
- Herbst RS, Heymach JV, Lippman SM (2008) Lung cancer. *N Engl J Med* 359(13):1367
- Shakeel PM, Burhanuddin MA, Desa MI (2022) Automatic lung cancer detection from CT image using improved deep neural network and ensemble classifier. *Neural Comput Appl* 34(15):1–14
- Braveen M, Nachiyappan S, Seetha R, Anusha K, Ahilan A, Prasanth A and Jeyam A (2023) ALBAE feature extraction based lung pneumonia and cancer classification. *Soft Comput* 1–14
- Sori WJ, Feng J, Godana AW, Liu S, Gelmecha DJ (2021) DFD-Net: lung cancer detection from denoised CT scan image using deep learning. *Front Comp Sci* 15:1–13
- Rendon-Gonzalez E and Ponomaryov V (2016) Automatic Lung nodule segmentation and classification in CT images based on SVM. In: 2016 9th International Kharkiv Symposium on Physics and Engineering of Microwaves, Millimeter and Submillimeter Waves (MSMW), pp 1–4, IEEE
- Wahab Sait AR (2023) Lung cancer detection model using deep learning technique. *Appl Sci* 13(22):12510
- Makaju S, Prasad PWC, Alsadoon A, Singh AK, Elchouemi A (2018) Lung cancer detection using CT scan images. *Procedia Comput Sci* 125:107–114
- Pelosi P, Abreu MGD (2010) Lung CT scan. *Open Nucl Med J* 2(1):86–98
- Swensen SJ, Jett JR, Sloan JA, Midthun DE, Hartman TE, Sykes AM, Aughenbaugh GL, Zink FE, Hillman SL, Noetzel GR, Marks RS (2002) Screening for lung cancer with low-dose spiral computed tomography. *Am J Respir Crit Care Med* 165(4):508–513
- Xiuhua G, Tao S and Zhigang L (2011) Prediction models for malignant pulmonary nodules based-on texture features of CT image. In: Theory and Applications of CT Imaging and Analysis, IntechOpen
- Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, Van Der Laak JA, Van Ginneken B, Sánchez CI (2017) A survey on deep learning in medical image analysis. *Med Image Anal* 42:60–88
- Nazir I, Haq IU, Khan MM, Qureshi MB, Ullah H, Butt S (2021) Efficient pre-processing and segmentation for lung cancer detection using fused CT images. *Electronics* 11(1):34
- Singh GAP, Gupta PK (2019) Performance analysis of various machine learning-based approaches for detection and classification of lung cancer in humans. *Neural Comput Appl* 31:6863–6877
- Sajja T, Devarapalli R, Kalluri H (2019) Lung cancer detection based on CT scan images by using deep transfer learning. *Traitement du Signal* 36(4):339–344
- Wang Y, Chen Q, Zhang B (1999) Image enhancement based on equal area dualistic sub-image histogram equalization method. *IEEE Trans Consum Electron* 45(1):68–75
- Jena SR, George ST, Ponraj DN (2021) Lung cancer detection and classification with DGMM-RBCNN technique. *Neural Comput Appl* 33:15601–15617
- Lung cancer detection using CT images dataset is taken from, "<https://wiki.cancerimagingarchive.net/display/Public/LIDC-IDRI>", accessed on Dec 2023
- Maini R and Aggarwal H (2010) A comprehensive review of image enhancement techniques. arXiv preprint. [arXiv:1003.4053](https://arxiv.org/abs/1003.4053)
- Murugesan B, Sarveswaran K, Shankaranarayana SM, Ram K, Joseph J and Sivaprakasam M (2019) Psi-Net: shape and boundary aware joint multi-task deep network for medical image segmentation. In: 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp 7223–7226, IEEE
- Chinniah P, Maram B, Velrajkumar P, Vidyadhari C (2022) Deep-Joint segmentation-based lung segmentation and hybrid optimization-enabled deep learning for lung nodule classification. *Int J Pattern Recognit Artif Intell* 36:2252021
- Banerjee A, Das N and Santosh KC (2022) Weber local descriptor for image analysis and recognition: a survey. *The Visual Computer*, pp 1–23
- Liu L, Lao S, Fieguth PW, Guo Y, Wang X, Pietikäinen M (2016) Median robust extended local binary pattern for texture classification. *IEEE Trans Image Process* 25:1368–1381
- Suresh A, Shunmuganathan KL (2012) Image texture classification using gray level co-occurrence matrix based statistical features. *Eur J Sci Res* 75:591–597
- Han, D., Kim, J. and Kim, J., “Deep pyramidal residual networks”, In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 5927–5935, 2017.
- Feng L and Yang G (2022) Deep Kronecker Network. arXiv preprint arXiv: 2210.13327

30. Babu UR, Kumar VV, Sujatha B (2012) Texture classification based on texton features. *Int J Image Graph Signal Process* 4(8):36
31. Lessa V and Marengoni M (2016) Applying artificial neural network for the classification of breast cancer using infrared thermographic images. In: Computer Vision and Graphics: International Conference, ICCVG 2016, Warsaw, Poland, September 19–21, 2016, Proceedings Springer International Publishing, vol. 8, pp 429–438
32. Gorji MB, Mozaffar M, Heidenreich JN, Cao J, Mohr D (2020) On the potential of recurrent neural networks for modeling path dependent plasticity. *J Mech Phys Solids* 143:103972
33. Patel RK, Kashyap M (2022) Automated diagnosis of COVID stages from lung CT images using statistical features in 2-dimensional flexible analytic wavelet transform. *Biocybern Biomed Eng* 42(3):829–841
34. Patel RK, Kashyap M (2023) Automated screening of glaucoma stages from retinal fundus images using BPS and LBP based GLCM features. *Int J Imaging Syst Technol* 33(1):246–261
35. Patel RK, Kashyap M (2023) Automated diagnosis of COVID stages using texture-based Gabor features in variational mode decomposition from CT images. *Int J Imaging Syst Technol* 33(1):807–821
36. Patel RK, Kashyap M (2023) Machine learning- based lung disease diagnosis from CT images using Gabor features in Littlewood Paley empirical wavelet transform (LPEWT) and LLE. *Comput Methods Biomed Eng* 11(5):1762–1776
37. Patel RK, Kashyap M (2023) The study of various registration methods based on maximal stable extremal region and machine learning. *Comput Methods Biomed Eng* 11(6):2508–2515
38. Bushara AR, Vinod-Kumar RS, Kumar SS (2023) Classification of benign and malignancy in lung cancer using capsule networks with dynamic routing algorithm on computed tomography images. *J Artif Intell Technol.* <https://doi.org/10.37965/jait.2023.0218>
39. Kunkayab T, Bahrami Z, Zhang H, Liu Z, Hyde D (2024) A deep learning-based framework (Co-ReTr) for auto-segmentation of non-small cell lung cancer in computed tomography images. *J Appl Clin Med Phys* 25(3):14297
40. Eldho KJ, Nithyanand S (2024) Lung cancer detection and severity analysis with a 3D deep learning CNN model using CT-DICOM clinical dataset. *Indian J Sci Technol* 17(10):899–910

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.