Multimedia Tools and Applications (2023) 82:37573–37592
https://doi.org/10.1007/s11042-023-14893-1

LCD-Capsule Network for the Detection and Classification of Lung Cancer on Computed Tomography Images

Bushara A.R.[1] · **Vinod Kumar R.S.**[1] · **Kumar S.S.**[2]

Received: 19 August 2022 / Revised: 27 January 2023 / Accepted: 6 February 2023 /
Published online: 23 March 2023

Abstract

Lung cancer is the second most prominent cancer in men and women, and it is also the leading cause of cancer-related mortality.If lung cancer is diagnosed early, when it is minuscule and has not spread, it is preferable to be adequately treated. A non-invasive low-dose Computed Tomography (CT) scan can detect abnormal patches in the lungs that could be cancerous. It is proposed that machine learning and pattern classification be used to identify and categorize lung cancer from CT scans. Pattern classification algorithms like deep learning can categorize input data into various classes based on the input's characteristic features. A novel deep learning framework formulated by the encapsulation of a Convolutional Neural Network (CNN) and a Capsule Neural Network (CapsNet) called LCD-CapsNet, leveraging the capabilities of these networks to minimize vast amounts of data and achieve spatial invariance for lung cancer detection and classification using CT images is proposed. The primary objective of the proposed method is to create algorithms that would classify and examine images from a dataset to determine whether or not a patient had or posed a danger of developing lung cancer. The Lung Image Database Consortium (LIDC) datasets are utilized to assess this deep learning model, from which 4335 images were collected for the training and testing pipeline. The results demonstrated that LCD-CapsNet outperforms CapsNet, with an average Precision of 95 %, Recall of 94.5 %, F1-Score 94.5 %, Specificity 99.07 %, Area Under the Curve of 0.989, and Accuracy 94 % of benign and malignant data.

Vinod Kumar R.S. and Kumar S.S. are contributed equally to this work.

✉ Bushara A.R.
bushara.ar@gmail.com

Vinod Kumar R.S.
rsvinodkumar@niuniv.com

Kumar S.S.
kumarss@niuniv.com

[1] Department of Electronics and Communication Engineering, Noorul Islam Center for Higher Education, Kumarakoil, Kanyakumari, 629180, Tamil Nadu, India

[2] Department of Electronics and Instrumentation Engineering, Noorul Islam Center for Higher Education, Kumarakoil, Kanyakumari, 629180, Tamil Nadu, India

Keywords Capsule network · Deep learning · Convolutional neural network · Computed tomography · Lung cancer

1 Introduction

The world has advanced technologically, yet cancer continues as a chronic disease, with lung cancer being the second most frequent [50]. This is because lung cancer symptoms emerge at the end of the disease's progression, and the patient is no longer able to be rescued. As a result, a person suffering from lung cancer can only be saved if cancer is diagnosed in its initial phases. Because of this, many methodologies were developed as well as computer-aided innovations [54], which are outstanding in their own right. Medical image analysis provides a significant benefit in the realm of health. Different imaging modalities may be used to identify lung cancer [13], and CT [36] is one of the most cost-effective, non-invasive, and efficient methods used in clinics to assess lung health [15]. Manual assessment of lung abnormalities predicted on CT scans is a a laborious effort that necessitates the services of a pulmonologist with extensive knowledge. If a computer-assisted approach is used to diagnose a lung abnormality, the evaluation report can be sent to a pulmonary specialist to aid in deciding on a procedure and selecting a course of treatment.

As a result, beginning in the early 1980s, doctors began to use Computer-Aided Diagnostic (CAD) [27] systems to aid in the interpretation of medical images [18]. Such systems investigate images using processes such as preprocessing, segmentation [37], object detection [2, 7], feature extraction, and classification [3]. Extracting features is a crucial part in a CAD framework of machine learning. Lung cancer has been diagnosed using image processing methods in the past [29]. By analyzing CT images obtained with artificial intelligence algorithms, machine learning techniques[1] assist in the diagnosis and classification of lung nodules. In today's fast-paced world of Artificial Intelligence (AI), image processing analytics utilizing Machine Learning [22, 38]and DL [14, 25] technology is highly effective for the accurate diagnosis and detection of lung cancer. DL has evolved tremendously since the advent of the Convolutional Neural Network (CNN) [10, 41] and it has been in use for more than two decades, providing an improved diagnosis.

In traditional CNNs, a certain cost function is utilized in order to measure the global error that builds up at the rear of the network as the training moves forward. On the other hand, if the weight among two neurons is judged to be zero, the activation of a neuron will not develop any further and will cease increasing at that point. This is because the activation of a neuron is dependent on the weight between two neurons. In the process known as recurrent dynamic routing, in addition to the agreement, the informtion is routed in accordance with the feature specifications. This is done in place of supplying a single size together with the comprehensive cost function.

The problems of vanishing gradient and saturated accuracy can arise if the CNN is trained to an excessive depth. Objects or sub-parts of objects can be represented by capsules, which are collections of neurons whose activity vectors can be used to extract structured features while preserving knowledge of their spatial relationships. An approach used in capsule networks, known as dynamic routing, makes use of the level of agreement among the input capsules to create the final product capsules. For dynamic routing, primary capsules are added together and their weights are adjusted iteratively during the forward based on whether closely their activations match the sum of all the weights. State-of-the-art accuracy can be attained through the use of this routing-by-agreement method.

To solve the aforementioned problems, this paper proposes a Lung Cancer Detection capsule network (LCD-CapsNet) for the diagnosis and categorization of lung cancer. Extracting spatial and temporal features from CT images is facilitated by the capsule network's superior capacity for learning the intrinsic relationship among various features. Accordingly, the capsule network was preferred as the foundational model for lung cancer classification rather than conventional CNNs. The summary of the contributions of this paper are,

(i) A novel LCD-CapsNet is structured to extract multi-dimensional CT image features to increase classification accuracy and classify lung cancer as either benign or malignant and,
(ii) The ideation and creation of an effective, highly provable, deep learning neural network based on capsules that is capable of achieving state-of-the-art results on CT images. The proposed LCD-CapsNet combines CNN and CapsNet to outperform state-of-the-art methods on small datasets.

The rest of the manuscript is organized as follows. Section 2 discussed the related works for the lung cancer classification. Section 3 explains the proposed methodology for lung cancer detection. The Capsule Networks and the proposed LCD-CapsNet are also discussed in Section 3. Section 4 presents the dataset that was used to evaluate the proposed LCD-CapsNet, as well as the experiments and findings. Section 5 concludes the paper.

2 Related works

In the field of deep learning, CNNs are a type of multi-layer neural network. Input and output layers, as well as hidden layers, constitute its structure. Most frequently, CNN is applied to images. The CNN will extract the pertinent image features through supervised training. The CNN weights are then determined using the back propagation algorithm version. The hidden layer typically comprises of several multiple stages.

The state-of-the-art CNN architectures [16] are used for image analysis and have produced several satisfying outcomes. It has also made significant progress in lung cancer detection and classification. Song et al. [51] constructed a CNN structure for the detection and classification of lung nodules and yielded an accuracy of 84.15 %. AlexNet is a robust CNN architecture for lung cancer diagnosis, and Hamdalla F. Al-Yasriy et al. [9] examined 110 test CT scans and found it to be 93.54 % accurate. Heng Yu et al. [57] devised a framework for determining lung cancer that uses a deep neural network after segmentation and has comparable accuracy. Abhir Bhandary et al. [15] suggested another deep learning network, a modified AlexNet, for lung cancer classification using the LIDC dataset. For the detection of lung nodules using CT images, H. Polat et al. [42] proposed a CNN model with and P. Wu et al. [55] presented a deep residual network. [33] proposed a 3D lung cancer detection method based on the Active Contour Model, as well as classification using various machine learning algorithms. With the aid of various feature extraction models, Asuntha et al.[12] suggested a deep learning model to detect the cancerous lung nodule.

T. Ghani [20] proposed regression as a machine learning approach for predicting lung cancer survival rates using 2D CT data. An adaptive boosting method was utilized by S. Pang et al. [39] to increase classification performance on a more efficient state-of-the-art CNN architecture, DenseNet, which achieved competent image classification results. A Generative model for lung cancer was established by Salama et al. by using ResNet [45]. G. E. Hinton et al. [44] developed Capsule Network (CapsNet) to alleviate the inadequacies

of CNN, and P. Afshar et al. [5] proposed MIXCAPS network for malignancy prediction in lung nodules based on CapsNet structure. H. J. D. Koresh et al. [26] segmented the major boundaries of the corneal layer and K. Adu et al. [4] proposed a novel method for lung and colon cancer classification based on Capsule Network. Fast CapsNet was used by Aryan Mobiny et al. [31] to suggest a technique for lung screening that was effective for lung nodules. Utilizing the Capsule network, B. Tang et al. [53] used pathological images to detect lung cancer.

Saleh et al. [47] suggested using a CNN in conjunction with the SVM algorithm for the classification of lung cancer. With the assistance of the data from chest CT scans, the CT images of the lungs are classified using a method that is a hybrid of CNN and SVM as either normal, adenocarcinoma, or large cell carcinoma. Squamous cell carcinoma is also a possibility. Kalaivani et al. [24] classified CT scans of the lungs as either normal or malignant using an adaptive boosting algorithm that was implemented using DenseNet Architecture that was trained on Clinical Data. Employing GoogleNet by Huseiny et al. [8] to identify images of malignant nodules found on computed tomography of the lung. The method involves submitting the input images to a pre-processing step that is straightforward and quick. This step extracts regions of interest from the images. The implementation of Harris Hawks optimization techniques, along with the assistance of Cancer Imaging Archive datasets, is used by Guo et al. [21] to improve CNN classifier results. The feature-based classifier was made use of the LBP and Haralick features, which are both relatively well-known. 3D Res U-Net is employed for segmentation, whereas 3D ResNet50 is utilized for the classification of lung cancer by Yu et al. [56] leveraging LIDC-IDRI.

Researchers are also utilizing Capsule Networks for the detection and classification of lung cancer. The concept of a memory-enhanced capsule network that uses LUNA-16 to extract features from high-dimensionmal input has been proposed by Mobiny et al.[32]. Afshar et al. [6] utilized LIDC-IDRI datasets in the process of implementing a 3D-MCN for lung nodule cancer. The 3D inputs in this method provide information about the 3D nodules, while the multi-scale inputs record the local characteristics. A design based on CapsNet that can handle only a few training samples. Shafi et al. [48] are accountable for the implementation of the Capsule Network with SVM using LIDC-IDRI datasets. Capsule Network is utilized in the process of segmenting lung nodules, while Support Vector Machines are utilized in the process of classifying lung cancer. Ramana et al.[43] utilized LIDC-IDRI datasets to put their Capsule Network and Whale optimization algorithm into action for lung cancer classification. The integration of capsule networks reduces the computational complexity as well as overfitting. The Whale optimization algorithm modifies the pre-trained network's hyperparameters to achieve a higher accuracy rate.The qualitative comparison of earlier methods with the recently proposed method for lung cancer classification leveraging deep neural network is presented in Table 1.

The authors reviewed deep learning-based CT image lung cancer detection and classification models. To improve accuracy, CNN training necessitates a substantial amount of data. CNNs are also sensitive to affine transformation and cannot maintain spatial relationships between image features. The authors use CT scan images to classify lung cancer using the powerful CapsNet. It increases all the performance criteria because the information contained in low-level capsules can be successfully used to determine the poses and the probabilities of high-level capsules, effectively saving a significant number of computational resources while also significantly enhancing the accuracy of image classification. The authors concluded by discussing the difficulties of developing the Capsule Network-based lung cancer detection system.

Table 1 Comparison of previous methods for Lung Cancer Classification

State-of-the-art Model	Methodology	Benchmarks	Key Points
Saleh et al. [47]	CNN+SVM	Chest CT-Scan images Dataset	lung CT images are classified using a hybrid CNN-SVM method into adenocarcinoma, large cell carcinoma, normal, or squamous cell carcinoma.
Kalaivani et al. [24]	DenseNet	Clinical Data	An adaptive boosting algorithm is used with DenseNet Architecture for lung cancer classification
Huseiny et al. [8]	GoogLeNet	IQ OTH/NCCD	Input images are fed to a simple and rapid pre-processing step that extracts regions of interest
Guo et al. [21]	CNN	Cancer Imaging Archive	Improved Harris Hawks optimization improved CNN classifier results. The feature-based classifier used two well-known Haralick and LBP features
Yu et al. [56]	3D ResNet50	LIDC-IDRI	For segmentation, 3D Res U-Net is used, while 3D ResNet50 is used for classification
Mobiny et al.[32]	Capsule Network	LUNA - 16	memory-enhanced capsule network extract features from high-dimensional input
Mobiny et al.[32]	3D-MCN	LIDC-IDRI	3D inputs provide 3D nodule information, while multi-scale input captures local features.
Shafi et al. [48]	Capsule Network+SVM	LIDC-IDRI	capsule Network is used to segment lung nodules while SVM is used to classify lung cancer
Ramana et al.[43]	Capsule Network+WOA	LIDC-IDRI	Whale optimization algorithm tunes pre-trained network hyperparameters to enhance the accuracy rate

The objective of this article is to develop a DL Architecture for early lung cancer diagnosis and classification using LCD-CapsNet which utilizes the characteristics of both Capsule Network and Convolutional Neural Network. CapsNet is a popular image categorization and detection algorithm. Despite the fact that CapsNet lacks a pooling operation, it still provides good classification accuracy with small datasets. This paper proposes a novel framework for extracting features from CT lung images and classifying them as benign or malignant. Few quantifying parameters namely, Accuracy, Precision, F1-Score, and Recall of the proposed method are also compared with state-of-the-art classifiers.

3 Methodology

The framework of the proposed LCD-CapsNet is shown in Fig. 1. The Inputs are pairs of images labeled as benign or malignant. This dataset is divided into three sections namely train, validation, and test data. The train and validation data are pre-processed before being fed into the feature extractor [23]. In the proposed model, CNN is used as a feature extractor which can extract the deep features of images. These image representative features are fed into the proposed LCD-CapsNet model. In LCD-CapsNet, CNN is utilized to extract the features, and capsule networks are employed to perform classification. Accuracy and Loss in training and validation are performed during model training. The accuracy of a proposed model test can be computed with the help of reserve test data. The predicted label is obtained using the softmax classifier. The classification accuracy is calculated using the measure of quality for both the predicted and true labels.

3.1 Datasets

The LIDC-IDRI Dataset contains 1,018 scans from 1,010 patients and 244,527 images. This dataset contains CT scan images and XML files with case-specific CT scan annotations. Four expert thoracic radiologists annotated these images. Radiologists first classified CT findings by sizes such as nodules less than 3 mm, nodules between 3 and 6 mm, and

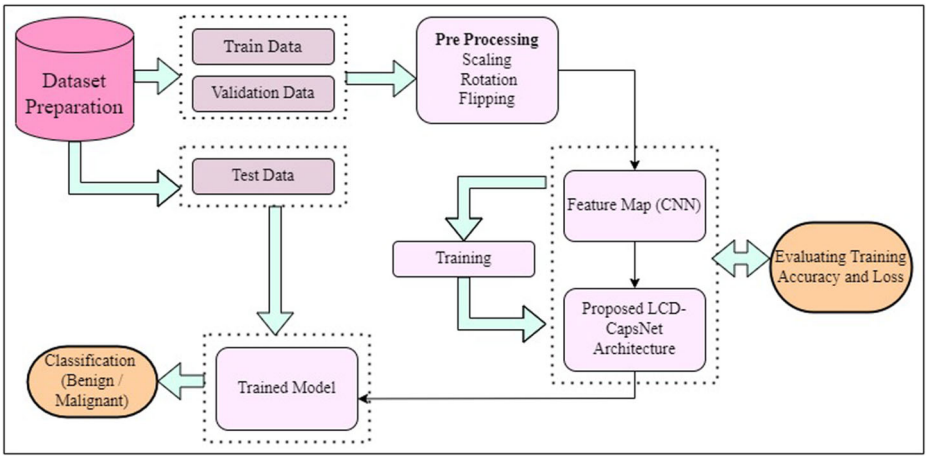

Fig. 1 The Work Flow of the Proposed LCD-CapsNet Model

non-nodules. The second phase involves a covert comparison of radiologist categorizations. Thus, all four radiologists evaluate each dataset nodule independently. This dataset enables nodule and two-label patient diagnosis triage. Each CT scan produces 64 to 764 DICOM images with 512 × 512 resolution. Unknown, Benign/Non-Cancer, Primary Lung Cancer/Malignant, and Metastatic Lesions are the LIDC-IDRI dataset's four lung nodule categories (Other Primary Cancer). Figure 4 shows LIDC-IDRI datasets of benign and malignant data.

3.2 Preprocessing

Pre-processing images reduces network load and computation complexity. An LIDC-IDRI CSV file contains lung CT scan images with labels. Using CSV annotations, the model separates and labels cancerous and noncancerous images. ImageDataGenerator from the Keras library performed data augmentation techniques like rescaling, rotation, and horizontal and vertical flips. Images can be rotated ten degrees from zero. Set range = 1 and use Rescale by 1./255 to scale each pixel value from [0, 255] to [0, 1]. The color mode function of the flow directory of Python can be used to convert grayscale images into color ones. The Images are downsized from 512 × 512 to 299 × 299, and the Keras ImageDataGenerator class enhances images efficiently and quickly.

3.3 Convolutional neural network

CNN has the advantage of less pre-processing than other classification algorithms. CNN retrieves image features [46] using pooling and convolution. These extracted features then train the model and classify images. The proposed model has four convolutional and one pooling layer as shown in Fig. 2. The convolutional layer filters the image in 3 × 3 pixels. The actual outcome of the convolutional layer is a matrix that is significantly more compact than the initial image. Backpropagation helps train the network by adding nonlinearity to convolutional layer input. Selecting the maximum or minimum value from each group of values in the pooling layer reduces the activation layer matrix size. Finally, Capsule Network classifies CNN output.

3.4 Capsule network

Capsule Networks, also known as CapsNet, are a type of Artificial Neural Network (ANN) [34] that are designed to overcome the information loss that occurs during the pooling

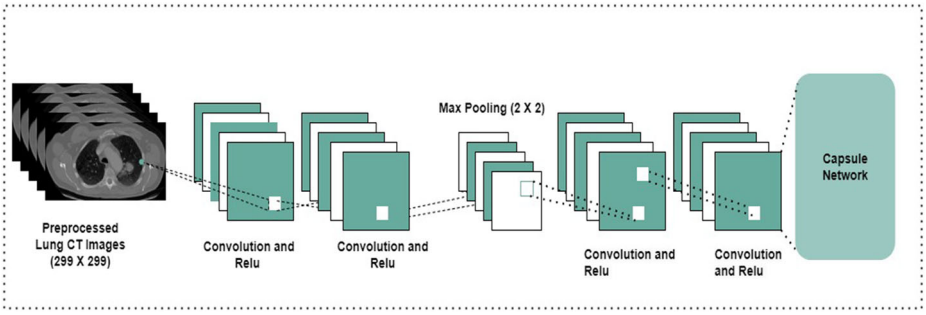

Fig. 2 Block Diagram of Proposed CNN for Feature Extraction

operations carried out by CNN. These networks are able to retrieve spatial information as well as other crucial characteristics. A supervised neural network receives an image as input and outputs a label. The equivariance structure is preferred for tasks that need location-related information, such as segmentation [35] and detection since it preserves the position and its pose across the network. CNN pooling layers downsample the image, which causes the network to gradually lose location-related information. The potential problems of CNN were addressed by Hinton [44] and implemented CapsNet in 2017 as shown in Fig. 3. In a Capsule network, output of each neuron is no longer a scalar, but rather a vector that represents the entity. The angle of the vector represents the entity's properties, and the vector length goes from one to near zero if the entity hardly exists. Otherwise, the closer a vector's length is to one, the more likely it is to exist. The lower layer of the capsule network is connected to the higher layer via a matrix that predicts the whole object based on the recognized part. The weighted sums of all the prediction vectors in a parent capsule are used to calculate the final result vector.

During training, the network employs the routing by agreement technique, in which each prediction vector is a dot product with the output vector, with the results gradually adjusting the connection weights. Let the layers p and $p+1$ ought to have m and n capsules respectively. Calculating the activations of the capsules at layer $p+1$ is the objective given the activations at layer p. Initially, the capsules at layer p are used to compute the prediction vectors. The prediction vector $\hat{v}_{j|i}$ for j of layer $p+1$ by a capsule i of the layer, p is provided by (1).

$$\hat{v}_{j|i} = W_{i,j} v_i \tag{1}$$

where $W_{i,j}$ is the weight matrix and v_i is capsule activations at layer t. The weighted total of all the prediction vectors f_j generated by the capsules of layer t for the capsule j is then used to compute the output vector for the capsule j as shown in (2).

$$f_j = \sum_i C_{i,j} \hat{v}_{j|i} \tag{2}$$

Between capsules i and j, the scalar $C_{i,j}$ stands for a Coupling Coefficient. The capsules of layers p and $p+1$ are connected by different coupling coefficients $C_{i,j}$, which will be adjusted depending on the image content. It can be established by a dynamic routing algorithm as in (3).

$$C_{i,j} = \frac{e^{b_{i,j}}}{\sum_k e^{b_{i,k}}} \tag{3}$$

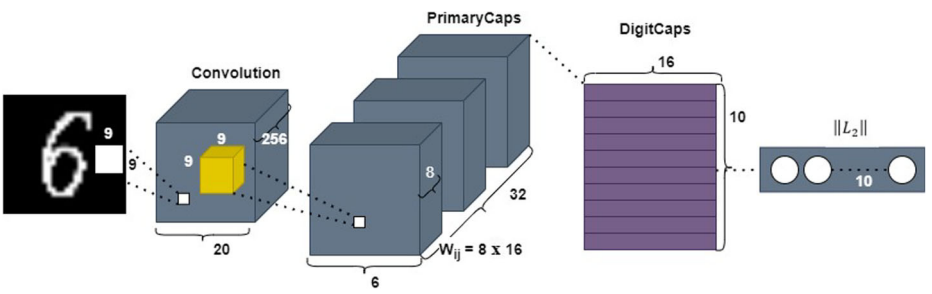

Fig. 3 Basic Capsule Network Model

The output vector f_j is squashed to get the activation s_j of the capsul e j. The squashing function is applied to make the probability of the magnitude of the output vector between 0 and 1. At the layer $p+1$, for a capsule j, $s_j = squash(f_j)$ and is calculated using (4).

$$Squashfunction, s_j = \frac{\|f_j\|^2 . f_j}{1 + \|f_j\|^2 . \|f_j\|} \tag{4}$$

where s_j represents the capsule's vector output and f_j represents the capsule's final input. When $\|f_j\| = 0$, the derivative of $\|f_j\|$ is unbounded, and it has the potential to explode during training. Calculating the square root of the sum of squares as well as a small epsilon number ϵ, provides the solution as in (5).

$$\|f_j\| = \sqrt{\sum_i s^2 + \epsilon} \tag{5}$$

Finally, the scalar product a_{ij} is used by (6), to calculate the agreement.

$$a_{i,j} = f_j . \hat{v}_{j|i} \tag{6}$$

During the entire process, the Coupling Coefficient $c_{i,j}$ learns from the input data, whereas the weight matrix $W_{i,j}$ adapts and modifies itself during training in CapsNet based on back propagation.

3.5 LCD-CapsNet

3.5.1 Overview of the architecture

The architecture of the LCD-CapsNet is shown in Fig. 4. The overview of LCD-CapsNet is as follows. There are four convolutional layers and two capsule layers in this network. Input to this is pre-processed lung CT images.

During pre-processing, 512 × 512 images were sampled down to 299 × 299 images, to shorten training time and reduce the number of parameters.In addition, each pixel value undergoes a rescaling transformation, shifting it from [0,255] to [0,1], as well as a rotation and flipping (both horizontal and vertical) of the image.

The first layer is the convolution layer which learns 64 filters, convolves with a kernel of size 3 × 3 and stride 1, and produces a feature map of size 297 × 297 × 64. The batch

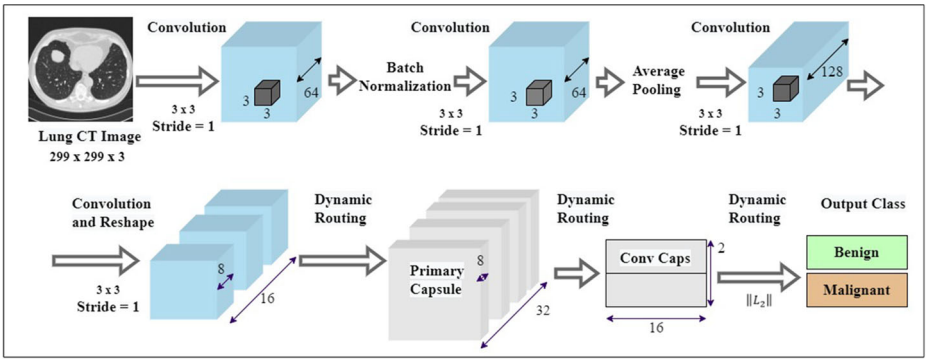

Fig. 4 LCD-CapsNet Architecture for Lung Cancer Detection and Classification

normalization used after the convolution layer, helps to maintain output mean zero and output variance one. The second layer is a convolution 295 × 295 × 64, followed by average pooling that reduces output shape to 147 × 147 × 64. The third and fourth convolution layers learn 128 filters and reshape the output of the fourth layer, which is the input of the Capsule Network. To activate the network, the ReLu function is used across all four convolutional layers.

The LCD-CapsNet contains two Capsule layers that execute the routing by the agreement method explained in Section 5.2. The primary capsule layer comprises 32 capsules of size 8 each. The final Capsule layer named Conv Caps includes the instantiation parameters for the 2 types with a dimension of 16 namely benign and malignant. The squash function is deployed as the non-linear activation function in the capsule network, with the length of these two capsules representing the probability of each category occurring.

3.5.2 Routing by agreement algorithm

Dynamic routing involves transmission of lower capsules data to a higher capsule, which is more appropriate capsule. It is the parent capsule that receives the results from child capsules. Using the agreement and assignment mechanism, which is based on the dot product, expectation-maximization, and mixture models, parent capsules carry out routing. The capsule with the greatest dot product is selected as the parent capsule. A dot product is calculated by multiplying the weight matrix by the prediction vector that was generated by the lower capsule layers.

The classification is formed by computing the length $\|L2\|$ of the capsule vectors in the final layer and allocating the malignant class to those capsule vectors whose magnitude are greater than the threshold, and the benign class to those capsule vectors whose magnitude are less than the threshold. The schematic representation in Fig. 5 explains dynamic routing algorithm. The input vector v_i, also known as the lower-level capsule is subjected to an affine transformation. The prediction vector $\hat{v}_i$ is obtained by multiplying each of these vectors v_i by the weight vector W_{ij}. The weighted sum of these prediction vectors with the coupling coefficient C_{ij}, which is learned through routing by agreement is computed. The

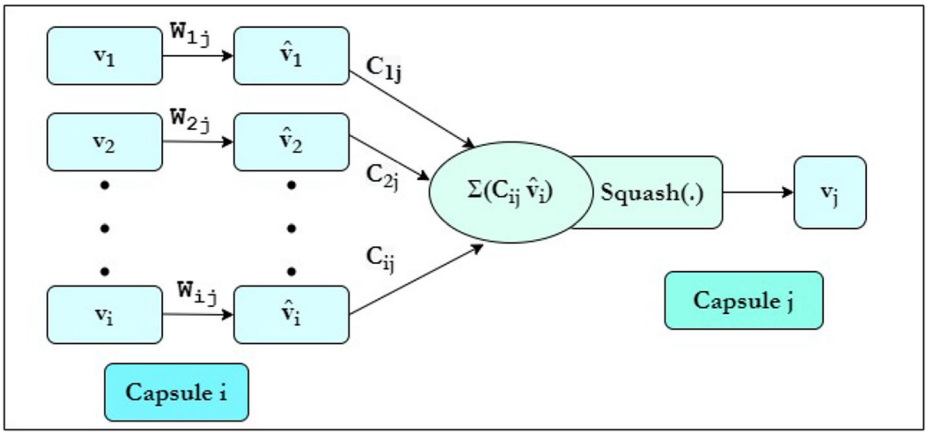

Fig. 5 Schematic Representation of Dynamic Routing Algorithm

weighted sum runs through a non-linearity squash function to ensure that the vector length is between 0 and 1. Algorithm 1 summarizes the pseudocode for LCD-CapsNet agreement routing.

procedure Routing ($\hat{v}_{j|i}, z, p$)
for in layer p and capsule j in layer $(p+1)$, and for every capsule $i : b_{i,j} \leftarrow 0$.
for z iterations do
for in layer p, all capsules $i : ci \leftarrow softmax(bi)$: using Eq. (3)
for in layer $(p+1)$ all the capsules j :
$f_i \leftarrow \sum_i c_{ij} \hat{v}_{j|i}$
for in layer $(p+1)$ all the capsules j:
$s_j \leftarrow squash(f_j)$; using Eq. (4)
for in layer p and capsule j in layer $(p+1)$ all the capsules i: $b_{i,j} \leftarrow b_{i,j} + f_j.\hat{v}_{j|i}$
return f_i

Algorithm 1 Routing by Agreement between Capsules

3.5.3 Loss function

The loss function is utilized to assess the accuracy of the network's predictions. Equation (7) is used to compute the loss in the LCD-CapsNet:

$$L_m = T_m \max(0, h^+ - \| f_m \|)^2 + \lambda(1 - T_m) max(0, \| f_m \| - h^-)^2 \tag{7}$$

where L_m - Loss term for one Conv Caps, T_m - Loss function of Conv Caps,

λ is numerical stability coefficient whose value is fixed at 0.5. $T_m \max(0, h^+ - \| f_m \|)^2$, calculates the loss for accurate ConvCaps, when T_m is 1. $\lambda(1 - T_m) max(0, \| f_m \| - h^-)^2$ computes the loss for erroneous ConvCaps, when T_m is 0.

4 Experimental results and discussion

The proposed system was evaluated using the LIDC dataset [11]. A total of 4335 CT images were utilized in this experiment for evaluation. Of these 2160 are benign and 2175 are malignant data. Four radiologists examined the LIDC data set in depth. This dataset depicts the size of nodules present in the lung using CT scans. If the size of the nodule is less than 3mm, it is deemed benign and if the size of the nodule is larger than 3mm, it is labeled as malignant. Figure 6 shows the samples of benign and malignant data of LIDC datasets. The datasets are broken down into three categories namely training, testing, and validation sets. From the LIDC dataset, 20% of the images (867) are used as test data. Of the remaining 80% of the image, 20% (694) are used for validation which is used to identify the best performing model, and the leftover datasets (2774) are considered as training sets for training the model.

With a learning rate of 10^{-3}, 50 epochs, a fixed batch size of 16, and this system was trained using an Adam optimizer. The non-linear activation function employed in Capsule is the squashing function. It reduces the length of a large vector to near one and the length of a small vector to zero. The experiments were carried out using Python libraries such as Keras and TensorFlow on a GPU NVIDIA-MX450, an i5-1135G7 processor, with Windows as the operating system.

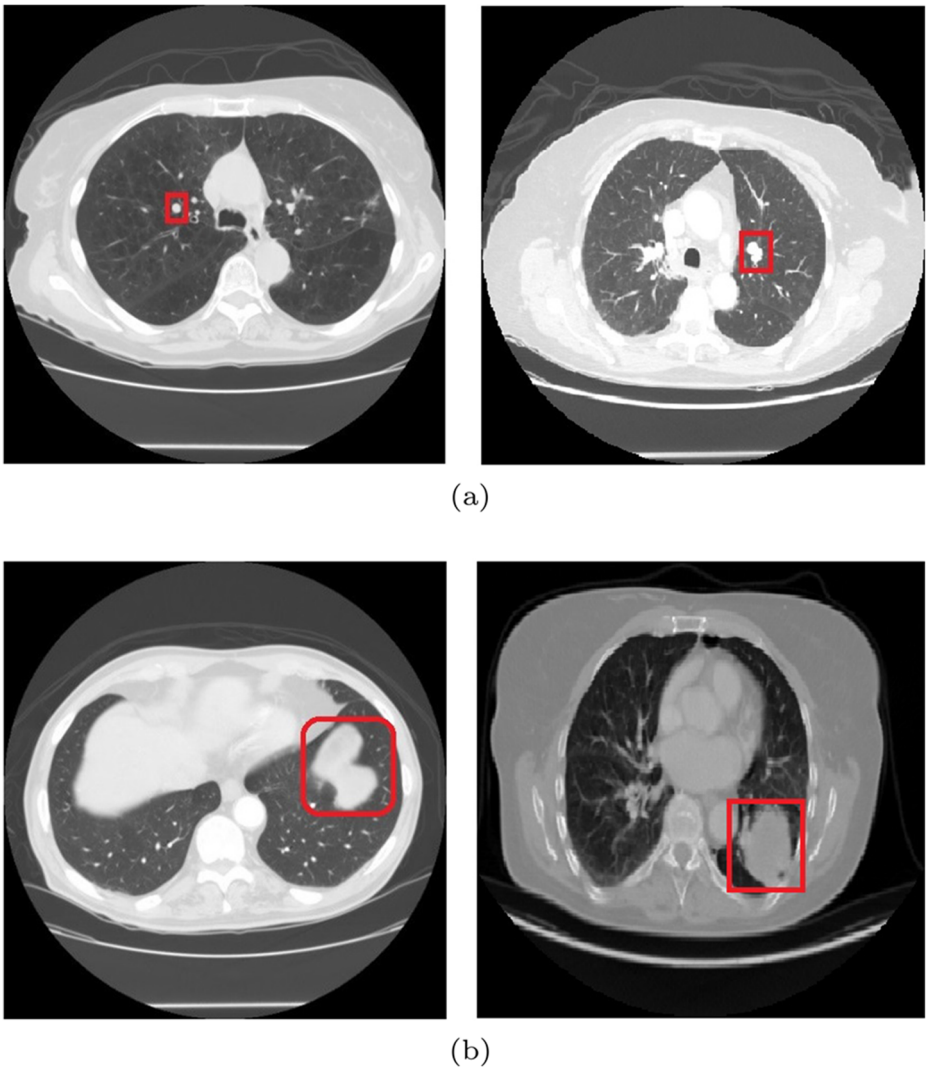

Fig. 6 Sample Lung Images (a) Benign nodule (b) Malignant nodule (Courtesy: LIDC- IDRI Dataset)

Two experiments have been done in this paper. They are detection and classification of Lung cancer using CapsNet, and LCD-CapsNet. In CapsNet, two secondary capsule layers were used after the primary capsule, and the total number of parameters were 295,488. But in the proposed LCD-CapsNet, only one secondary capsule layer is used, therefore the number of parameters could be reduced to 293,440. Feature map in Fig. 7 visualizes an inside story of LCD-CapsNet identifying different features present in lung images, thereby providing a deeper understanding of the algorithm of the proposed model. The abstract features are learned while moving deeper inside the neural network. In the initial stages of a neural network, it is possible to interpret the details that the filter is attempting to extract from the images. However, as the network progresses deeper, the process becomes

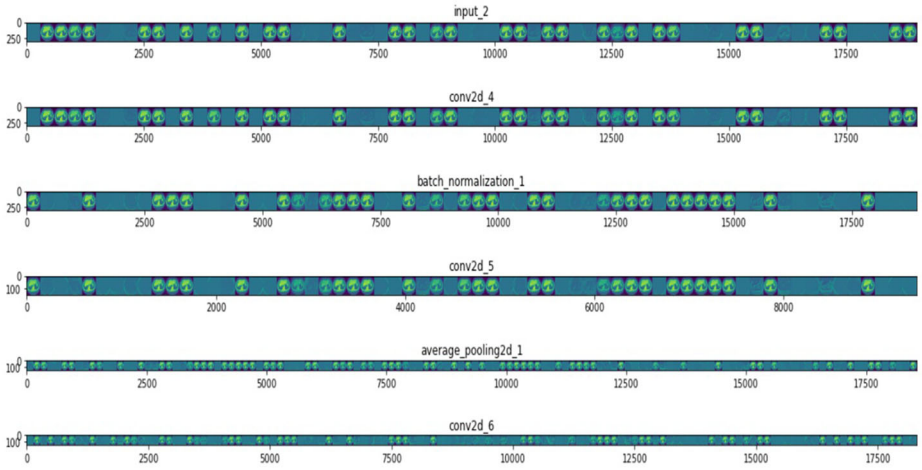

Fig. 7 Feature map for Lung image of LCD-CapsNet

more complex. The developed system was successfully trained and tested. The performance evaluation matrices of accuracy, recall, precision, and F1 score were also determined.

The amount of time necessary to train the models is listed in Table 2, together with the number of network parameters that required to be learned during the training procedure (expressed on the scale of millions). It is clear that the training and testing of the LCD-CapsNet is significantly faster than that of the conventional CapsNet for the data sets.

Figure 8 represents the training progress of CapsNet and LCD-CapsNet in which, Fig. 8 a) depicts the accuracy curve of training and validation of CapsNet, whereas Fig. 8 b) illustrates the training and validation accuracy of LCD-CapsNet. By evaluating these accuracy curves, it is known that the LCD-CapsNet provides improved accuracy than CapsNet. Figure 8c) and d) illustrate the training and validation loss of CapsNet and LCD-CapsNet respectively. The loss was reduced to 0.048 in CapsNet after 50 epochs, whereas the final loss in the proposed LCD-CapsNet is 0.037. Table 3 illustrates the confusion matrices, which indicate the effectiveness of the categorization models CapsNet and LCD-CapsNet.

There are 376 True Positive (TP) instances in CapsNet where the system predicted the existence of malignancy and 428 True Negative (TN) cases where the system predicted the absence of cancer in CapsNet. In addition, 4 False Positive (FP) instances projected malignancy but did not have cancer, indicating a Type 1 error, while 59 False Negative (FN) cases predicted benign but were malignant, indicating a Type 2 error. However, in the LCD-CapsNet system, the instances of TP, TN, FP, and FN are 392, 427, 5, and 43, respectively. From the comparison matrices shows in Table 1 it is understood that LCD-CapsNet provides better classification than CapsNet.

Table 2 Number of parameters and Processing time Comparison of CapsNet and LCD-CapsNet

Method	Parameters	Processing Time (s)
CapsNet	0.295M	1457
LCD-CapsNet	0.293M	1304

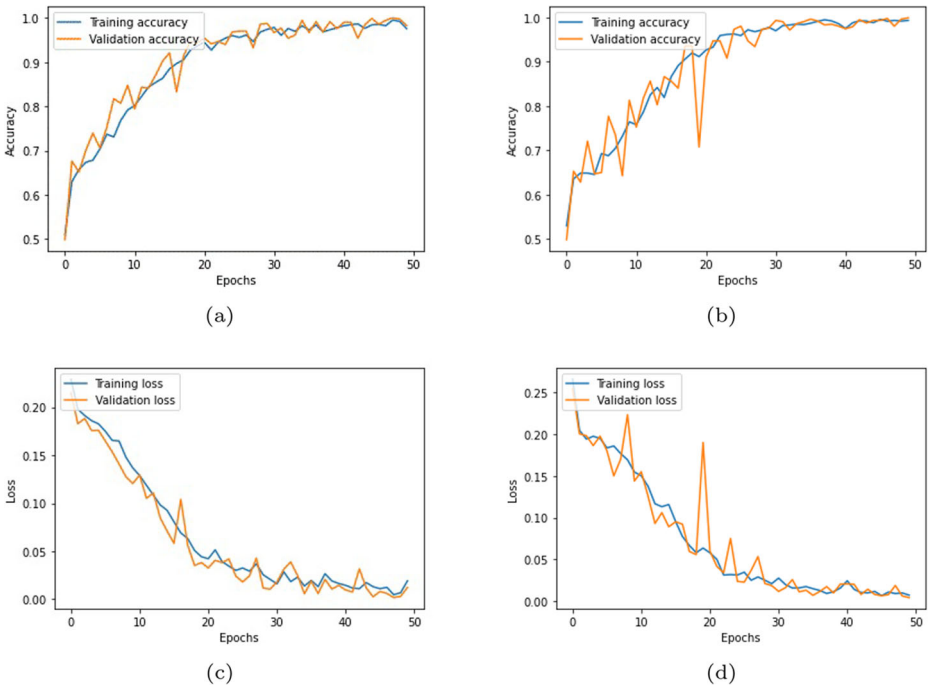

Fig. 8 Training progress of CapsNet and proposed LCD-CapsNet (a) Accuracy of CapsNet (b) Accuracy of LCD-CapsNet (c) Loss of CapsNet (d) Loss of LCD-CapsNet

The performance analysis matrices for the imposed CapsNet and LCD-CapsNet are shown in Tables 4. The inferences from the performance comparison are as follows. In LCD-CapsNet, all the performance metrics have been improved, as shown in Tables 4. Out of 867 test samples, 432 are benign and 435 are malignant. In CapsNet, the value of Recall or sensitivity and F1-Score is 92.5 %, whereas it is 94.5 % in LCD-CapsNet. Precision has increased from 93 % to 95 % and test accuracy has increased from 93 % to 94 % .

Table 5 compares the proposed LCD-CapsNet model with the existing cutting-edge CNN lung cancer classification methods. As shown in Table 5, LCD-CapsNet outperformed other state-of-the-art architectures in terms of accuracy, AUC, sensitivity, and specificity. Over [5], its accuracy is increased by 3.3%, AUC increased by 0.033, sensitivity increased by 4.5%, and specificity increased by 5.67%.

Table 3 Confusion Matrices of CapsNet and LCD-CapsNet

Model	Actual Class	Predicted Class	
		Benign	Malignant
CapsNet	Benign	428	4
	Malignant	59	376
LCD-CapsNet	Benign	427	5
	Malignant	43	392

Table 4 Performance analysis matrices for the proposed CapsNet and LCD-CapsNet

Quantifying Metrices	Model	
	CapsNet	LCD-CapsNet
Parameters	295,488	293,440
Precision (%)	93.5	95.0
F1-Score (%)	92.5	94.5
Recall(Sensitivity) (%)	92.5	94.5
Specificity (%)	98.5	99.1
Accuracy (%)	93.0	94.0

The performance of the proposed system is further visualized by the ROC curve and AUC value as shown in Fig. 9. The ROC curve is sketched with the sensitivity on the true positive rate on the y-axis and the 1- specificity (False Positive Rate) on the x-axis. While CapsNet has an AUC of 0.985, LCD-CapsNet has an AUC of 0.989 which is evident from Fig. 9. From these discussions it is understood that LCD-CapsNet outperforms CapsNet on LIDC datasets for lung cancer classification.

The baseline CapsNet model even outperformed the rest of the models in terms of accuracy, AUC, sensitivity, and specificity. Sensitivity and specificity are two important metrics used to gauge a test's accuracy and reliability. In this case, sensitivity is defined as the ability to detect cancer. It is critical to avoid providing a false negative.

A patient's clinical picture will be altered if a false negative is detected. Also, specificity refers to a test's ability to avoid false positives. False-positive results are undesirable for a cancer test because they may give the patient the false impression that they have been exposed to cancer in the past. LCD-CapsNet has better sensitivity, accuracy, and specificity than the state-of-the-art models, which is known from Fig. 10. In Fig. 10, x-axis represents the percentage value of accuracy, sensitivity and specificity and y-axis represents the methods used.

Table 6 presents a comparison between the recently suggested LCD-CapsNet model and the already deployed machine learning lung cancer detection methods. In terms of accuracy, sensitivity, and specificity, LCD-CapsNet performed significantly better than other methods, as shown in Table 6.

Table 5 Comparison of LCD-CapsNet with classical CNN methods

Literature	Accuracy(%)	AUC	Sensitivity(%)	Specificity(%)
Sori et al. [52]	87.80	-	87.40	89.10
Shen et al. [49]	87.14	0.930	77.00	90.03
Song et al. [51]	84.15	-	83.96	84.34
Causey et al. [17]	87.90	0.938	87.90	87.90
Afshar et al. [5]	90.70	0.956	89.50	93.40
Baseline	93.00	0.985	92.50	98.50
Proposed Work	94.00	0.989	94.50	99.07

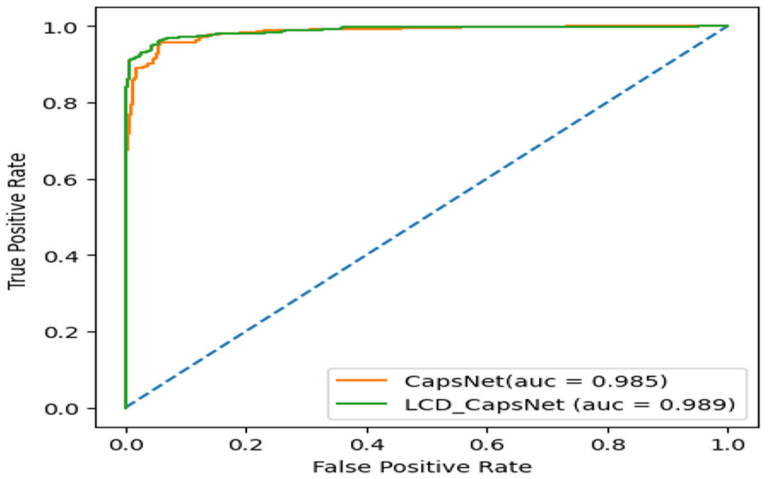

Fig. 9 Comparison of ROC of the Proposed LCD-Capsnet with CapsNet

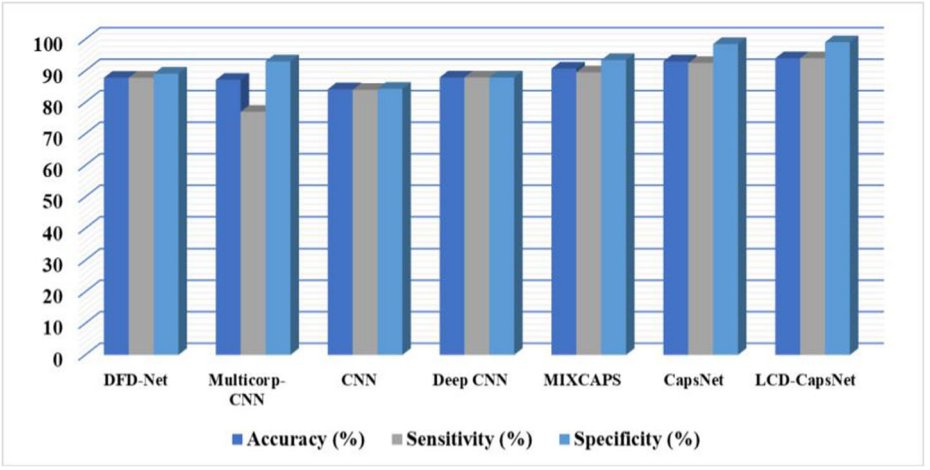

Fig. 10 Comparison of the proposed LCD-CapsNet method with the existing methods in terms of specificity, sensitivity, and accuracy

Table 6 Comparison of LCD-CapsNet with Machine Learning Methods

Literature	Model	Accuracy(%)	Sensitivity(%)	Specificity(%)
Parveen et al.[40]	SVM	-	91.38	89.56
Lennartz et al.[28]	KNN	87.00	94.00	67.00
Askary et al.[19]	Random Forest	90.73	90.67	90.8
Manju et al. [30]	SVM	87.00	88.00	-
Proposed	LCD-CapsNet	94.00	94.50	99.07

The Area Under the ROC Curve (AUC) is a numerical representation of the diagnostic test's overall performance. Fig. 11 depicts the AUC comparison of the proposed LCD-CapsNet model with other methods, demonstrating that the presented model is superior to the others.

Although the proposed LCD-Capsnet has achieved the highest accuracy, it does have some drawbacks. The most significant of which are that its algorithm is sluggish, primarily as a result of the inner loop of the dynamic routing algorithm, and that its implementation is more difficult than that of CNNs.

5 Conclusion and recommendation

This paper proposes two deep learning frameworks for Lung Cancer prediction based on LCD-CapsNet and CapsNet to enhance the performance in diagnosing Lung Cancer using CT images. The proposed LCD-CapsNet Architecture was developed using algorithms that would classify and analyze images from LIDC IDRI datasets to diagnose whether a patient even had lung tumor or was at possibility of acquiring it. A convolutional neural network was deployed to extract high-level features, and a capsule neural network was utilized to classify them. Lung nodule malignancy and benignness prediction in LIDC-IDRI datasets are compared.

The cross-validation results are quite promising, showing that Capsule Networks are better than CNN models at gathering pose information and spatial correlations, as well as distinguishing between malignant and benign images. Eventhough the amount of trainable parameters in LCD-CapsNet are fewer than in CapsNet, the results showed that LCD-CapsNet has excellent performance with limited set. The findings show that LCD-CapsNet is superior to CapsNet, with an average Accuracy of 94 % , Precision of 95 %, Recall of 94.5 %, F1-Score of 94.5 %, Specificity of 99.07 %, and AUC of 0.989 for both benign and malignant data. The objective of the study is fulfilled by achieving better results of performance matrices while using a smaller sample size of image datasets.

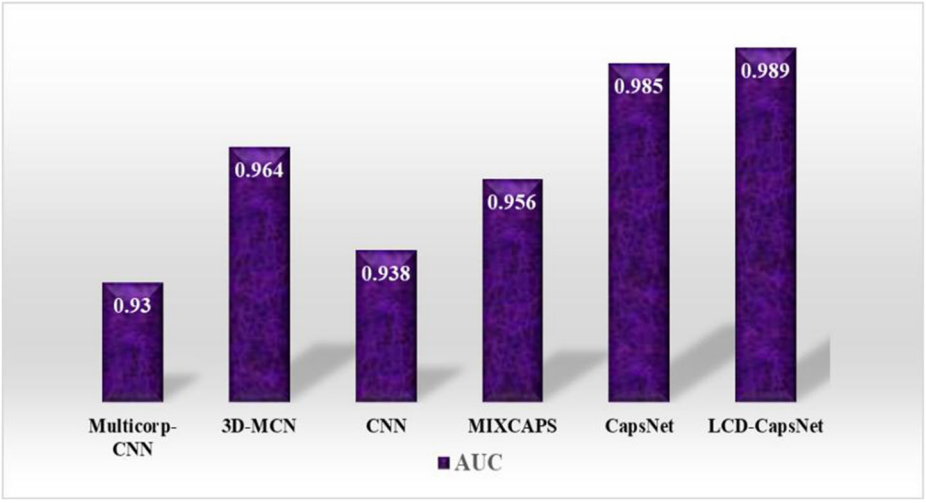

Fig. 11 AUC Comparison of the proposed LCD-CapsNet model with the existing methods

In the future research, VGG can be utilized as an alternative to CNN in the LCD-CapsNet input, to analyze an extension of CapsNet by increasing the number of capsule layers to alleviate the constraints of CapsNet.

Acknowledgment The authors acknowledge the National Cancer Institute and the National Institutes of Health for their contributions to the development of the LIDC/IDRI database, which is free and open to the public. There was no specific grant awarded for this study by any government, commercial, or nonprofit organization.

Data Availability LIDC-IDRI data are publicly availabe

Declarations

Conflict of Interests The authors declare that they have no conflict of interest.

References

1. Abubeker KM, Baskar S (2022) A Machine Learning Strategy for Internet-of-Things-Enabled Diabetic Prediction to Mitigate Pneumonia Risk. In: In 2022 10th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO) (pp. 1-6). IEEE
2. Abubeker KM, Baskar S (2022) Wireless sensor and wireless body area network assisted biosensor network for effective monitoring and prevention of non-ventilator hospital-acquired pneumonia. Frontiers in Sustainable Cities 4
3. Adebisi OA, Ajagbe SA, Ojo JA, Oladipupo MA (2022) Computer techniques for medical image classification: a review. Intelligent Healthcare, pp 19–36
4. Adu K, Yu Y, Cai J, Owusu-Agyemang K, Twumasi BA, Wang X (2021) DHS-Capsnet: Dual horizontal squash capsule networks for lung and colon cancer classification from whole slide histopathological images. Int J Imaging Syst Technol 31(4):2075–2092
5. Afshar P, Naderkhani F, Oikonomou A, Rafiee MJ, Mohammadi A, Plataniotis KN (2021) MIXCAPS: A capsule network-based mixture of experts for lung nodule malignancy prediction. Pattern Recognit 116:107942
6. Afshar P et al (2020) 3D-MCN: A 3D Multi-scale Capsule Network for Lung Nodule Malignancy Prediction. Sci Rep 10(1):7948. https://doi.org/10.1038/s41598-020-64824-5
7. Ajagbe SA, Oki OA, Oladipupo MA, Nwanakwaugwum A (2022) Investigating the efficiency of deep learning models in bioinspired object detection. In: 2022 International Conference on Electrical, Computer and Energy Technologies (ICECET) (pp. 1-6). IEEE
8. Al-Huseiny MS, Sajit AS (2021) Transfer learning with GoogLeNet for detection of lung cancer. Indones J Electr Eng Comput Sci 22(2):1078. https://doi.org/10.11591/ijeecs.v22.i2.pp1078-1086
9. Al-Yasriy HF, Al-Husieny MS, Mohsen FY, Khalil EA, Hassan ZS (2020) Diagnosis of lung cancer based on CT scans using CNN. IOP conf. Ser Mater Sci Eng 928(2)
10. Alzubaidi L, Zhang J, Humaidi AJ, Al-Dujaili A, Ye D, Al-Shamma O, Santamaría J, Fadhel MA, Al-Amidie M, Farhan L (2021) Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. J Big Data 8:53
11. Armato SG et al (2011) The lung image database consortium (LIDC) and image database resource initiative (IDRI): A completed reference database of lung nodules on CT scans. Med Phys 38(2):915–931
12. Asuntha A, Srinivasan (2020) A Deep learning for lung Cancer detection and classification. Multimed Tools Appl 79:7731–7762
13. Avanzo M, Stancanello J, Pirrone G, Sartor G (2020) Radiomics and deep learning in lung cancer. Strahlentherapie und Onkol 196(10):879–887
14. Aydın N, Çelik Ö, Aslan AF, Odabaş A, Dündar E, Şahin MC (2021) Detection of lung cancer on computed tomography using artificial intelligence applications developed by deep learning methods and the contribution of deep learning to the classification of lung carcinoma. Current Medical Imaging 17(9):1137–1141
15. Bhandary A, Ananth Prabhu G, Rajinikanth V, Palani Thanaraj K, Satapathy SC, Robbins DE, Shasky Charles, Zhang Yu-Dong, Tavares JMRS, Sri Madhava Raja N (2020) Deep-learning framework to detect lung abnormality – A study with chest X-Ray and lung CT scan images. Pattern Recognit Lett 129:271–278

16. Bushara AR, Vinod Kumar RS (2022) A deep learning-based lung cancer classification of CT images using augmented convolutional neural networks, Electronic Letters on Computer Vision and Image Analysis 21(1). https://doi.org/10.5565/rev/elcvia.1490
17. Causey JL, Zhang J, Ma S et al (2018) Highly accurate model for prediction of lung nodule malignancy with CT scans. Sci Rep 8(1):1–12
18. Doi K (2007) Computer-aided diagnosis in medical imaging: historical review, current status and future potential. Comput Med Imaging Graph 31(4-5):198–211
19. El-Askary NS, Salem MAM, Roushdy MI (2019) Feature extraction and analysis for lung nodule classification using random forest. In: Inproceedings of the 2019 8th International Conference on Software and Information Engineering, pp 248–252
20. Ghani T, Oommen JB (2021) On utilizing 2D features from 3D scans to enhance the prediction of lung cancer survival rates. Pattern Recognit Lett 152:56–62
21. Guo Z, Xu L, Si Y, Razmjooy N (2021) Novel computer-aided lung cancer detection based on convolutional neural network-based and feature-based classifiers using metaheuristics. Int J Imaging Syst Technol 31(4):1954–1969. https://doi.org/10.1002/ima.22608
22. Huidrom R, Chanu YJ, Singh KM (2022) Neuro-evolutional based computer aided detection system on computed tomography for the early detection of lung cancer. Multimed Tools Appl 81:32661–32673
23. Jain S, Salau AO (2019) An image feature selection approach for dimensionality reduction based on kNN and SVM for AkT proteins. Cogent Eng 6(1):1599537
24. Kalaivani N, Manimaran N, Sophia S, Devi DD (2020) Deep learning based lung cancer detection and classification, IOP conf. Ser Mater Sci Eng 994(1). https://doi.org/10.1088/1757-899X/994/1/012026
25. Khan A, Sohail A, Zahoora U, Qureshi AS (2020) A survey of the recent architectures of deep convolutional neural networks. Artif Intell Rev 53(8):5455–5516
26. Koresh HJD, Chacko S, Periyanayagi M (2021) A modified capsule network algorithm for oct corneal image segmentation. Pattern Recognit Lett 143:104–112
27. Kumar SS, Nithila EE (2016) Lung cancer diagnosis from CT images using CAD system: a review. Int J Biomed Eng Technol 21(4):311–321
28. Lennartz S et al (2021) Texture analysis of iodine maps and conventional images for k-nearest neighbor classification of benign and metastatic lung nodules. Cancer Imaging 21(1):1–10
29. Makaju S, Prasad PWC, Alsadoon A, Singh AK, Elchouemi A (2018) Lung cancer detection using CT scan Images. Procedia Computer Science 125:1877–0509
30. Manju BR, Athira V, Rajendran A (2021) Efficient multi-level lung cancer prediction model using support vector machine classifier. In: IOP Conference Series: Materials Science and Engineering (Vol. 1012, No. 1, p. 012034). IOP Publishing
31. Mobiny A, Van Nguyen H (2018) Fast CapsNet for Lung Cancer Screening. In: Frangi A., Schnabel J., Davatzikos C., Alberola-López C., Fichtinger G. (eds) Medical Image Computing and Computer Assisted Intervention – MICCAI 2018. MICCAI 2018. Lecture Notes in Computer Science, vol 11071. Springer, New York
32. Mobiny A et al (2021) Memory-Augmented Capsule network for adaptable lung nodule classification. IEEE Trans Med Imaging:1–11. https://doi.org/10.1109/TMI.2021.3051089
33. Naqi SM, Sharif M, Lali IU (2019) A 3D nodule candidate detection method supported by hybrid features to reduce false positives in lung nodule detection. Multimed Tools Appl 78:26287–26311
34. Nasser IM, Abu-Naser SS (2019) Lung cancer detection using artificial neural network. International Journal of Engineering and Information Systems (IJEAIS) 3(3):17–23
35. Nithila EE, Kumar SS (2016) Segmentation Of lung nodule in CT data using active contour model and Fuzzy C-mean clustering. Alex Eng J 55(3):2583–2588
36. Nithila EE, Kumar SS (2017) Automatic detection of solitary pulmonary nodules using swarm intelligence optimized neural networks on CT images. Eng Sci Technol Int J 20(3):1192–202
37. Nithila EE, Kumar S (2019) Segmentation of lung from CT using various active contour models. Biomed Signal Process Control 47:57–62
38. Olalekan A, Jain S (2021) Informatics in Medicine Unlocked Adaptive diagnostic machine learning technique for classification of cell decisions for AKT protein. Informatics Med Unlocked 23:100511. https://doi.org/10.1016/j.imu.2021.100511
39. Pang S, Zhang Y, Ding M, Wang X, Xie X (2020) A deep model for lung cancer type identification by densely connected convolutional networks and adaptive boosting. IEEE Access 8:4799–4805
40. Parveen SS, Kavitha C (2014) Classification of lung cancer nodules using SVM Kernels. Int J Comput Appl 95(25)
41. Phani Kumar M, Dutta S, Murmu NC (2021) Tool wear classification based on machined surface images using convolution neural networks. Sādhanā 46:130

42. Polat H, Mehr HD (2019) Classification of pulmonary CT images by using hybrid 3D-deep convolutional neural network architecture. Appl Sci 9(5)
43. Ramana K, Kumar MR, Sreenivasulu K, Gadekallu TR (2022) Early prediction of lung cancers using deep saliency capsule and Pre-Trained deep learning frameworks. Front Oncol 12(June):1–13. 10.3389/fonc.2022.886739
44. Sabour S, Frosst N, Hinton GE (2017) Dynamic routing between capsules. Adv Neural Inf Process Syst,pp 3856–3866
45. Salama WM, Shokry A, Aly MHA (2022) Generalized framework for lung Cancer classification based on deep generative models. Multimed Tools Appl 81:32705–32722
46. Salau AO, Min M (2019) Feature extraction: A survey of the types, techniques, applications. In: 5th IEEE International Conference on Signal Processing and Communication (ICSC), Noida, India, pp158–164. https://doi.org/10.1109/ICSC45622.2019.8938371
47. Saleh AY, Chin CK, Penshie V, Al-Absi HRH (2021) Lung cancer medical images classification using hybrid cnn-svm. Int J Adv Intell Informatics 7(2):151–162. https://doi.org/10.26555/ijain.v7i2.317
48. Shafi I et al (2022) An effective method for lung cancer diagnosis from CT scan using deep Learning-Based support vector network. Cancers (Basel) 14(5457):1–18
49. Shen W, Zhou M, Yang F, Dongdong Y, Di D, Yang C, Zang Y, Tian J (2017) Multi-crop Convolutional Neural Networks for lung nodule malignancy suspiciousness classification. Pattern Recognit 61:663–673
50. Siegel RL, Miller KD, Fuchs HE, Jemal A (2021) Cancer statistics. CA Cancer J Clin 71(1):7–33
51. Song QZ, Zhao L, Luo XK, Dou XC (2017) Using deep learning for classification of lung nodules on computed tomography images. J Healthcare Eng 2017(8314740):7
52. 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 Comput Sci 15(2)
53. Tang B, Li A, Li B, Wang M (2019) Capsurv: Capsule network for survival analysis with whole slide pathological images. IEEE Access 7:26022–26030
54. Thakur SK, Singh DP, Choudhary J (2020) mLung cancer identification: a review on detection and classification. Cancer Metastasis Rev 39(3):989–998
55. Wu P, Sun X, Zhao Z, Wang H, Pan S, Schuller B (2020) Classification of lung nodules based on deep residual networks and migration learning. Comput Intell Neurosci, vol 2020
56. Yu H, Li J, Zhang L, Cao Y, Yu X, Sun J (2021) Design of lung nodules segmentation and recognition algorithm based on deep learning. BMC Bioinforma 22(314):1–20. https://doi.org/10.1186/s12859-021-04234-0
57. Yu H, Zhou Z, Wang Q (2020) Deep learning assisted predict of lung cancer on computed tomography images using the adaptive hierarchical heuristic mathematical model". IEEE Access 8:86400–86410