



CAD system for lung nodule detection using deep learning with CNN

R. Manickavasagam¹ · S. Selvan² · Mary Selvan³

Received: 6 November 2020 / Accepted: 29 September 2021 / Published online: 22 November 2021
© International Federation for Medical and Biological Engineering 2021

Abstract

The early detection of pulmonary nodules using computer-aided diagnosis (CAD) systems is very essential in reducing mortality rates of lung cancer. In this paper, we propose a new deep learning approach to improve the classification accuracy of pulmonary nodules in computed tomography (CT) images. Our proposed CNN-5CL (convolutional neural network with 5 convolutional layers) approach uses an 11-layer convolutional neural network (with 5 convolutional layers) for automatic feature extraction and classification. The proposed method is evaluated using LIDC/IDRI images. The proposed method is implemented in the Python platform, and the performance is evaluated with metrics such as accuracy, sensitivity, specificity, and receiver operating characteristics (ROC). The results show that the proposed method achieves accuracy, sensitivity, specificity, and area under the roc curve (AUC) of 98.88%, 99.62%, 93.73%, and 0.928, respectively. The proposed approach outperforms various other methods such as Naïve Bayes, K-nearest neighbor, support vector machine, adaptive neuro fuzzy inference system methods, and also other deep learning-based approaches.

Keywords Lung nodules · Computed tomography (CT) · Deep learning · Convolutional neural network (CNN) · LIDC

1 Introduction

Lung cancer is one of the most serious health problems in the world. The mortality rate of lung cancer is the highest among all other types of cancer. Almost 80–85% of lung cancer belongs to non-small cell type lung cancer (NSCLC) [1]. However, the survival rate of lung cancer patients can be increased if nodules are diagnosed accurately.

Lung nodule detection is a time-consuming and complex process, and radiologists must carefully analyze the nodules in CT scan images. Doi [2] shows that 30% of lung nodules overlay with the other common anatomic structures,

which may lead to the missing of the nodules by radiologists. The lung images are obtained through several imaging techniques. Among these, the CT image is a basic imaging method for the detection of lung nodules used by many researchers.

A number of methods have been developed for lung nodule detection and classification. Gong et al. [3] have proposed a probabilistic-based Naïve Bayes classifier which is used for effective binary classification. The performance of this method is improved over Fisher's linear discriminant analysis (FLDA) in terms of accuracy and sensitivity. The K-nearest neighbor (K-NN) method, a probabilistic-based method suitable for 2-class classification, has been proposed by Mao et al. [4]. This method provides a classification accuracy of 89% which is better than methods used for lung nodule detection such as random forest, regression tree, and learning vector quantization. Support vector machine (SVM)-based method, a powerful method for classification problems, has been proposed for lung nodule classification by Han et al. [5]. The rule-based filtering process combined with features of the region of interest (ROI) is used to reduce the false-positive (FP) rate. The adaptive neuro fuzzy inference system (ANFIS) combines the merits of ANN (artificial neural network) and FIS (fuzzy inference system) to get better performance of lung nodule detection. Tariq et al. [6]

✉ R. Manickavasagam
manick6apr1979@gmail.com

S. Selvan
drselvan@gmail.com

Mary Selvan
ms.maryselvan@gmail.com

¹ Department of BME, Alpha College of Engineering, Chennai-124, India

² Department of ECE, St. Peter's Institute of Higher Education and Research, Avadi, Chennai-54, India

³ Department of CSE, St. Peter's College of Engineering and Technology, Chennai-54, India

proposed a neuro fuzzy classifier approach which provides accurate and effective detection of lung nodules.

Recently, deep learning-based methods are extensively used to provide effective solutions for many applications such as natural language processing, speech recognition, and image analysis. The advantage of deep learning in CAD systems is that it can perform nodule detection by learning with the automatically extracted features during training. The deep learning architectures consist of different types of neural networks with increased hidden layers when compared with traditional machine learning processes. Convolutional neural network (CNN) is widely used for object recognition and classification with a remarkable increase in accuracy [7, 8].

The CNN architecture consists of convolution layers to perform the convolution operation to attain an improved understanding of input images for effective classification. Here, every neuron is associated with others such that it reacts to the receptive field surrounding it. In addition to this, the learnable parameters are extracted with convolution and pooling processes in CNN. As a result, the CNN approach produces improved accuracy for image classification. Many CNN-based approaches are proposed, including 7 layers Imagenet [9], 8 layers AlexNet [10], 25 layers VGG net [11], and 152 layers ResNet [12, 13], for various applications of pattern classification.

Kumar et al. [14] presented a deep learning technique with a stacked autoencoder and achieved an accuracy of 75.01% for lung nodule classification. Shen et al. [15] analyzed lung cancer detection using the LIDC database with a multiscale CNN method and achieved an accuracy of 86.84% for lung nodule detection. Shin et al. [16] evaluated the performance of transfer learning on lung CT images with CNN and attained 91.1% as nodule detection accuracy. Li et al. [17] proposed an 8-layer CNN-based deep learning method for the effective detection of lung nodules and achieved an accuracy of 92.4%. Jiang et al. [18] developed a deep learning model with CNN for computer-assisted diagnosis of lung cancer and achieved an accuracy of 94% with reduced false positives.

Since detection of lung nodules more accurately will lead to better and prompt treatment by oncologists, new approaches with improved performance are always needed. For improving the performance of the lung nodule detection

using CT images, a new CNN-based deep learning method, CNN-5CL, which consists of 11 layers, is proposed in this paper. This work focuses on developing a unified method for ROI segmentation of the lung region from the database image along with the classification of nodules. CNN constructed with the rectified linear unit (ReLU) function [19] is used for automatic feature extraction from the normalized lung region and classification of cancerous lung images with greater accuracy.

This introduction section briefs the need for this proposed method and related works. Section II describes the proposed CAD system for lung nodule detection. Section III provides the results obtained through the implementation. Section IV offers the discussion on the results of the proposed method. Section V provides the conclusion of the paper.

2 Proposed method

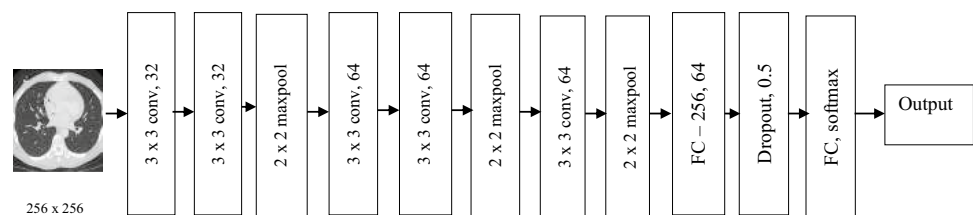
The proposed CNN-5CL architecture for the lung nodule classification consists of 11 layers. The classification of lung nodules is carried out using LIDC/IRDI database images [20, 21].

2.1 Convolutional neural networks (CNNs)

A convolutional neural network (CNN) is a multilayer neural network, which has many convolution layers and also has a required number of fully connected layers as a typical multilayer neural network. CNN deep learning methods are developed with the concepts of local perception, the sharing of weights, and sampling in terms of space or time. Local perception can detect maximum local features of the image as basic features for classification or analysis. Another important merit of CNN is that more features of the input can be used for training to make an efficient decision for classification [22].

The proposed CNN-5CL architecture is given in Fig. 1. CNN architecture is evaluated with a different number of layers, neurons, and filters, and the effects of these changes in the performance are discussed in Sections III and IV. The selected layers, filters, and neurons in each layer after the evaluation are given in Fig. 1. It is composed of 5 convolutional layers; each convolution layer is composed of a set of

Fig. 1 Proposed 2D CNN-5CL architecture



filters for convolution. These convolution operations create a set of feature maps. The number of feature maps is equal to the number of filters used for convolution. Every feature map is related with one convolution kernel (i.e., weight), and each convolution kernel represents a feature, such as the edge of the image. The size of input images used in this work is 256×256 in grayscale form.

It is always advantageous to use the max pooling layer after the convolution layers. Instead of using all the convolutional layers together, if the convolutional layers are used in small numbers followed by a max pooling layer, the performance is improved [23]. This reduces the number of parameters used for processing. The 11 layers are split into 4 groups. The first and second groups have two convolutional layers and one max pooling layer. The third group contains one convolutional layer and one max pooling layer. The last group consists of a fully connected dense layer, dropout layer, and fully connected softmax layer.

2.2 Convolution layer

The input image is split into small sizes based on the convolutional kernel, and a convolution operation is carried out on the image. The convolution layer has sparse interactions with the input and kernel function of the filter. It can be seen in CNN architecture that the kernel size is smaller than the size of the input. The advantages of this sparse interaction are the detection of useful features such as edges, reduction in memory space, and accuracy of classification. The convolution function has a weight-sharing capability, and this results in the extraction of diverse features of an image and efficient parameter learning similar to that of fully connected networks. The output of the convolutional layer is obtained based on Eq. (1) [24].

$$(n \times n \times n_c) * (f \times f \times n_c) = (n - f + 1) * (n - f + 1) * n_c' \tag{1}$$

where $n \times n$ is the size of the input image, $f \times f$ is the size of the filter kernel, n_c is the number of channels ($n_c = 1$ for 2D grayscale image), and n_c' is the number of filters used for convolution.

The activation function used in the convolutional layer is ReLU (rectified linear unit) because of its effectiveness in computation. The ReLU activation function is given in Eq. (2) [25].

$$f(x) = \begin{cases} x, & \text{if } x \geq 0, \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

where x is the input and $f(x)$ is the activation function.

2.3 Pooling layer

The pooling layer is added in the CNN to find similar elements, and it is used to reduce the size of the pooling layer output image. The pooling layer is used to reduce the feature map of the specific position in convolution output. Three types of operations are involved in the pooling layer: min pooling, mean (average) pooling, and max pooling. Min pooling estimates the neighborhood within a minimum of feature points, average pooling estimates the average neighborhood within the feature points, and max pooling estimates the neighborhood within a maximum of feature points. The max pooling operation is chosen due to its ability to reduce the output error [3]. The max pooling operation is performed on the output image of the convolution layer as mentioned in Eq. (3) [26].

$$f_{\max}(x) = \max(x_i) \tag{3}$$

where x_i is the i th element of the max pooling output: $i = 1, 2, 3, 4$ for 2×2 kernel.

2.4 Fully connected layer

The fully connected layers along with the dropout layer are used for classification in the last section of CNN. The operation of the fully connected layer is global which is different from convolution and pooling layer operations. Therefore, this operation produces a nonlinear grouping of features which are used for classification. The fully connected layer is used to flatten the output of the max pooling layer in the form of an array as it represents the feature vector of the input from all the neurons of the previous layer.

2.5 Dropout layer

The dropout process is a practice of random selection of neurons from the previous layer for training. The dropout layer uses max-norm regularization, which improves the efficiency of classification through random skipping of connections with selected probability. Here, the dropout layer is used to reduce the number of parameters extracted through convolution and max pooling operations, and this layer can avoid the over fitting of features during the training process. Dropout rate represents the fraction of parameters considered from the available parameters. A proper dropout rate is to be chosen for achieving improved performance.

Finally, the fully connected softmax layer is used to identify the nodule availability based on the probability of the class label [27] given in Eq. (4).

$$S_i = \frac{e^{z_i}}{\sum_{j=1}^k e^{z_j}} \quad (4)$$

where S_i is the probability value of softmax output for class i ($i = 1, 2$), z_i is the softmax layer output value of neuron i ($i = 1, 2$), and k is the number of neurons in the output layer ($k = 2$).

2.6 Dataset

Images from the Lung Image Database Consortium (LIDC) [28] are used for evaluating the performance of the proposed system. The dataset constructed from the LIDC database consists of 1018 images of non-nodules and nodules with a minimum nodule diameter of 3 mm along with the annotations [29]. Among the 1018 images, 135 images did not have nodules, and the remaining 983 images have 2669 nodules. The input segmented database contains 5338 nodule images and 1796 non-nodule images with sizes of 256×256 . The datasets with segmented images are further divided into a training set (80%) and a testing set (20%) for the evaluation. The number of images used for training and testing is 5867 and 1467, respectively. The sample input images of the nodule and non-nodule categories are shown in Fig. 2.

2.7 Performance metrics

The metrics considered for performance evaluation are accuracy (correctly classifies nodule and non-nodule images), sensitivity (correctly determines the nodule images), and specificity (correctly determines the non-nodule images). The mathematical relationships of these parameters are provided in Eqs. (5) to (7).

$$\text{Accuracy} = \frac{\text{TN} + \text{TP}}{(\text{TN} + \text{TP} + \text{FN} + \text{FP})} \quad (5)$$

$$\text{Sensitivity} = \frac{\text{TP}}{(\text{TP} + \text{FN})} \quad (6)$$

$$\text{Specificity} = \frac{\text{TN}}{(\text{TN} + \text{FP})} \quad (7)$$

where TP (true positive) denotes the images are correctly classified as nodules; FP (false positive) denotes the images are wrongly classified as nodules; TN (true negative) denotes the images are correctly classified as non-nodules; and FN (false negative) denotes the images are wrongly classified as non-nodules.

Receiver operating characteristics (ROC), another way of evaluating the performance, is plotted between true positive rate (TPR) and false positive rate (FPR). This is used to find the overall performance of the classifier. From this, the area

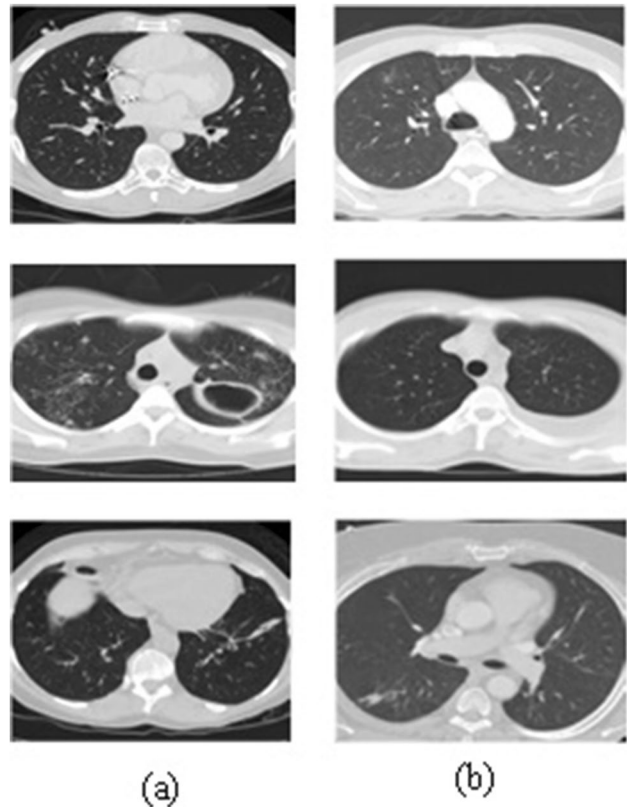


Fig. 2 Sample input images with a size of 256×256 : **a** nodule category and **b** non-nodule category

under the ROC curve (AUC) is calculated which represents the aggregate measure of performance across all possible classifications.

3 Results

The execution of the proposed CNN-5CL method is implemented in Python 3.7.4. The implementation platform uses Keras 2.1.3 with tensorflow as backend along with CPU acceleration. The experiments are conducted with different batch sizes of input. The batch size represents the number of input images that are used for training the network at a single instance.

The experiments are conducted with diverse batch sizes and convolutional layers, and the attained performances are listed in Table 1. With 128 batch size inputs and 5 convolutional layers, better performance is achieved.

Now the 5 convolutional layers are split into 3 groups, with two each in the first and second groups and one in the third group. All these groups will have one max pooling layer. The fourth group consists of a fully connected layer, dropout layer, and softmax layer.

Table 1 Performances of the CNN-5CL method for various number of convolution layers and input batch sizes

Input and network		Performance (%)		
Batch size	Convolution layers	Accuracy	Sensitivity	Specificity
128	4	84.02	97.84	61.93
	5	98.88	99.62	93.73
180	4	83.54	95.16	67.34
	5	96.29	98.36	84.73
256	4	82.24	93.11	56.15
	5	92.47	96.25	79.12

Table 2 Performances of the CNN-5CL method for various dropout rates

Dropout rate	Performance (%)		
	Accuracy	Sensitivity	Specificity
0.4	94.14	95.8	89.44
0.45	96.26	97.63	90.6
0.5	98.88	99.62	93.73
0.55	97.72	98.14	91.46
0.6	93.23	94.32	86.37

The selection of the number of filters, filter sizes, and the number of neurons in each layer is discussed in Section IV.

The performance of the proposed CNN method for various dropout rates is presented in Table 2. This analysis is carried out to select a suitable dropout rate to achieve better performance. From the results, the dropout rate of 0.5 is found to be suitable to attain the best performance.

With the proper selection of various components of the architecture, the layer specifications of the proposed CNN are given in Table 3.

Table 3 Layer specifications of the proposed model for 256 × 256 input

Group	Layer	Type	Input size	Kernel size	Number of filters	Output size	Number of parameters
1	1	Convolution	$256 \times 256 \times 1 = 65,536$	3×3	32	$254 \times 254 \times 32 = 2,064,512$	320
	2	Convolution	$254 \times 254 \times 32 = 2,064,512$	3×3	32	$252 \times 252 \times 32 = 2,032,128$	9246
	3	Max pooling	$252 \times 252 \times 32 = 2,032,128$	2×2	32	$126 \times 126 \times 32 = 1,016,064$	0
2	1	Convolution	$126 \times 126 \times 64 = 1,016,064$	3×3	64	$124 \times 124 \times 64 = 984,064$	18,496
	2	Convolution	$124 \times 124 \times 64 = 984,064$	3×3	64	$122 \times 122 \times 64 = 952,576$	36,928
	3	Max pooling	$122 \times 122 \times 64 = 952,576$	2×2	64	$61 \times 61 \times 64 = 238,144$	0
3	1	Convolution	$61 \times 61 \times 64 = 238,144$	3×3	64	$59 \times 59 \times 64 = 222,784$	36,928
	2	Max pooling	$59 \times 59 \times 64 = 222,784$	2×2	64	$29 \times 29 \times 64 = 53,824$	0
4	1	Fully connected	$29 \times 29 \times 64 = 53,824$	NA	64	53,824	409,856
	2	Dropout (dropout rate=0.5)	53,824	NA	-	26,912	204,928
	3	Softmax	26,912	NA	-	Output	Output

The ROC curve represents the performance of classification plotted between true positive rate (TPR) and false positive rate (FPR), where TPR is the same as sensitivity and FPR is 1-specificity. The ROC curve of the proposed method is shown in Fig. 3. The area under the ROC curve (AUC) value achieved is 0.928 which indicates the overall performance of the predictions.

The performance of the proposed CNN-5CL is compared with K-nearest neighbor (Mao et. al. [4]), Naïve Bayes (Gong et. al. [3]), SVM (Han et. al. [5]), ANFIS (Tariq et. al. [6]) methods, and other CNN approaches proposed by Li et. al. [17] and Jiang et. al.[18]. The results are shown in Table 4.

From the results shown in Table 4, it is seen that the proposed method outperforms other methods. The accuracy, sensitivity, specificity, and AUC of the proposed method are 98.88%, 99.62%, 93.73%, and 0.928, respectively, which are better than other machine learning and deep learning methods. Also, our method has considered all the images in the database LIDC/IDRI, whereas other methods considered a slightly lesser number of images for their experiments.

The comparison of computation time for training and testing of the proposed CNN-5CL method with other methods is shown in Table 5. It can be seen that the proposed method takes lesser time than the other methods for both training and testing. The computation time of 30.27 min and 19.23 s is achieved by our proposed method for training and testing, respectively.

4 Discussion

The architecture of the proposed CNN-5CL model is designed by conducting various experiments by varying the number of convolution layers, number of filters, and neurons in each layer and dropout in the dropout layer.

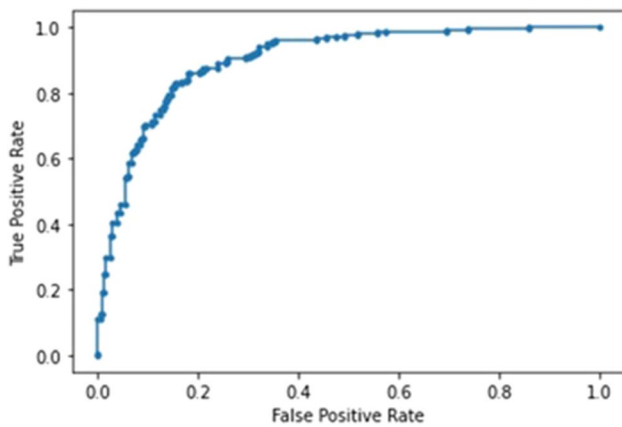


Fig. 3 ROC of the proposed method

Also, the batch size of the input for training is chosen to achieve the best performance.

The convolution layer consists of neurons which take the input of the selected area, called the receptive arena of the neuron, from the preceding layer. The area of the receptive arena of the neuron is square, and it is defined by the size of the filters used in convolution layers. Here, the different sizes of the filters including 3×3 , 5×5 , and 7×7 are used for evaluation. The sizes of the receptive arena of each neuron are 9, 25, and 49, respectively. It is observed that the performance is high with the filter size 3×3 when compared with filter sizes of 5×5 and 7×7 as it extracts local features of images effectively.

When the number of convolution layers is less, it leads to under fitting of the network parameters, and the performance is reduced. If the number of layers is more, it leads to the over fitting of the network parameters which affects the training process, and the performance is degraded. With a 3×3 filter size, experiments are conducted for 4 and 5 convolution layers. For 4 layers, 2 convolutional layers in group 1 and 1 convolution layer each in groups 2 and 3 is used for implementation. For 5 layers, the architecture shown in Fig. 1 is used. The results are shown in Table 1. Since the performance is better for 5 layers, this is selected. Hence,

2 convolution layers in the first and second groups and one convolution layer in the third group are selected as shown in Fig. 1.

Three max pooling layers, one for each group, are used in the model. When different sizes of filters such as 2×2 , 3×3 , and 4×4 are used for the evaluation, the filter size of 2×2 is found to be suitable as it preserves key features from the input. In the first, second, and third groups, one max pooling layer in each group is used. The number of neurons in each layer is the product of the size of the input and the number of filters in that layer.

The dropout rate is to be chosen in the dropout layer to achieve better performance. With different dropout rates (0.4, 0.45, 0.5, 0.55, and 0.6) the performance is studied. From Table 2, it can be seen that the performance measures such as accuracy, sensitivity, and specificity are better for the dropout rate of 0.5 when compared with other dropout rates. Hence, the dropout rate of 0.5 is selected.

After selecting the various layers, number of filters, filter sizes, and neurons, the architecture of the CNN-5CL model is proposed. Table 3 shows the layer specifications of the proposed model. This architecture is used for performance comparison with other methods.

Table 4 shows the comparison of performances of CNN-5CL with other methods. The Naïve Bayes (Gong et al. [3]) and K-nearest neighbor classifier (Mao et al. [4]) are probabilistic-based methods. Classification is carried out by these methods based on the configuration of feature-independent probability. These methods do not require a large size of samples for effective training. The SVM classifier (Han et al. [5]) is a powerful method for classification even for the larger dataset and high dimensional features. However, the selection of hyperplanes for classification is difficult. The ANFIS classifier (Tariq et al. [6]) gives good classification accuracy. However, the formation of rules for proper classification is very challenging. Deep learning methods proposed by Li et al. [17] and Jiang et al. [18] have produced better nodule detection accuracy. However, they have used only selected images for classification. Our method considers all the images in the LIDC database and offers improved classification results. Also, the automatic parameter tuning

Table 4 Comparison of performance of CNN-5CL with other methods

Method	Performance			
	Accuracy(%)	Sensitivity(%)	Specificity(%)	AUC
K-Nearest Neighbor (Mao et al. [15])	88.78	89.66	86.67	0.836
Naïve Bayes (Gong et al. [6])	93.48	93.33	93.75	0.862
SVM (Han et al. [8])	94.44	96.77	89.29	0.896
ANFIS (Tariq et al. [1])	96.67	98.38	92.86	0.913
CNN (Li et al. [21])	92.4	87.1	83.67	0.905
CNN (Jiang et al. [13])	94	95.8	89.16	0.912
CNN-5CL [Proposed]	98.88	99.62	93.73	0.928

Table 5 Comparison of computation time for training and testing of the proposed CNN-5CL method with other methods

Method	Overall computation time (CT) for training (min)	Average CT for one image for testing (s)
K-nearest neighbor (Mao et. al. [4])	57.32	55.28
Naïve Bayes (Gong et. al. [3])	53.43	42.56
SVM (Han et. al. [5])	48.53	38.42
ANFIS (Tariq et. al. [6])	41.26	26.54
CNN-5CL (proposed)	30.27	19.23

capability is achieved. The accuracy, sensitivity, specificity, and AUC of the proposed method are 98.88%, 99.62%, 93.73%, and 0.928, respectively.

The execution time taken for both training and testing is relatively less in the proposed method which is shown in Table 5. Automatic feature extraction reduces the execution time when compared with machine learning techniques. The proposed architecture requires less execution time when compared with other deep learning methods.

5 Conclusion

In this paper, the 11-layer CNN-5CL method is proposed for the effective classification of nodules and non-nodules from the LIDC/IDRI database of lung images. CNN-5CL architecture is proposed from the results of experiments conducted by varying the number of convolution layers, number of filters, number of neurons, and dropout rates. The performance of the proposed method is compared with the performance of other methods. The experimental results show that the proposed CNN-5CL method performs better when compared with K-nearest neighbor proposed by Mao et. al. [4], Naïve Bayes proposed by Gong et. al. [3], SVM proposed by Han et. al. [5], ANFIS method proposed by Tariq et. al. [6], and other CNN approaches proposed by Li et. al. [17] and Jiang et. al. [18]. The accuracy, sensitivity, specificity, and AUC of the proposed method are 98.88%, 99.62%, 93.73%, and 0.928, which are better than those of the other machine learning and deep learning methods. This method can be extended to develop and design high-performance CAD systems for the categorization of lung nodule types. Also, attempts can be made to apply this approach to other image formats of lung images for nodule classification.

Declarations

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

Conflict of interest The authors declare no competing interests.

References

- Liu J, Jiang H, Gao M (2017) An assisted diagnosis system for detection of early pulmonary nodule in computed tomography images. *J Med Syst* 41:30–43
- Doi K (2007) Computer-aided diagnosis in medical imaging: historical review, current status and future potential. *Comput Med Imaging Graph* 31:198–211
- Gong J, Liu J, Wang L, Zheng B, Nie S (2016) Computer-aided detection of pulmonary nodules using dynamic self-adaptive template matching and a FLDA classifier. *European Journal of Medical Physics* 2016:1502–1509
- Keming Mao, Zhuofu Deng, (2016) Lung nodule image classification based on local difference pattern and combined classifier, computational and mathematical methods in medicine, 2016: 1–7.
- Han H, Li L (2015) (2015) Fast and adaptive detection of pulmonary nodules in thoracic CT images using a hierarchical vector quantization scheme. In *Proceedings of IEEE International Journal of Biomedical and Health Informatics* 19(2):648–659
- Anam Tariq M, Usman Akram M, Javed Y (2013) Lung nodule detection in CT images using neuro fuzzy classifier. *IEEE Computational Intelligence in Medical Imaging* 2013:1–8
- Hinton GE, Osindero S, Teh YW (2006) A fast learning algorithm for deep belief nets. *Neural Comput* 18:1527–1554
- Song Q, Zhao L, Luo X, Dou X (2017) Using deep learning for classification of lung nodules on computed tomography images. *Journal of Healthcare Engineering* 2017:1–7
- Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition, *Proceedings of the IEEE International Conference on Learning Representations* 2014:1–14
- Krizhevsky A, Sutskever I, Hinton G. E (2012) ImageNet classification with deep convolutional neural networks, *Proceedings of the 25th International Conference on Neural Information Processing Systems*, 2012: 1097–1105.
- Hua KL, Hsu CH, Hidayati SC, Cheng WH, Chen YJ (2014) Computer-aided classification of lung nodules on computed tomography images via deep learning technique. *Onco Targets and Therapy* 2015:2015–2022
- Girshick R, Donahue J, Darrell T, Malik J (2014) Rich feature hierarchies for accurate object detection and semantic segmentation, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp.580–587.
- He K, Zhang X, Ren, S, Sun J (2016). Deep residual learning for image recognition, *Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016: 770–778.
- Kumar D, Wong A, Clausi D. A (2015) Lung nodule classification using deep features in CT images, *12th IEEE Conference on Computer and Robot Vision (CRV)*, 2015: 133–138.
- Shen W, Zhou M, Yang F, Yang C, Tian J (2015) Multi-scale convolutional neural networks for lung nodule classification, *Proceedings of the 24th International Conference on Information Processing in Medical Imaging*, pp. 588–599.
- Shin HC, Roth HR, Gao M (2016) Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset

- characteristics and transfer learning. *IEEE Trans Med Imaging* 35:1285–1298
17. Li W, Cao P, Zhao D, Wang J (2016) Pulmonary nodule classification with deep convolutional neural networks on computed tomography images. *Comput Math Methods Med* 2016:1–7
 18. Jiang H, Ma H, Qian W, Gao M, Li Y (2017) An automatic detection system of lung nodule based on multi-group patch based deep learning network. *IEEE Journal of Biomed Health Informatics* 22:1227–1237
 19. Choi WJ, Choi TS (2013) Automated pulmonary nodule detection system in computed tomography images: a hierarchical block classification approach. *Entropy* 15:507–523
 20. Kuruvilla J, Gunavathi K (2014) Lung cancer classification using neural networks for CT images. *Comput Methods Programs Biomed* 113:202–209
 21. LIDC-IDRI, (<https://wiki.cancerimagingarchive.net/display/Public/LIDC-IDRI>).
 22. LeCun Y, Kavukcuoglu K (2010) Farabet C (2010) Convolutional networks and applications in vision, *Proceedings of. IEEE International Symposium on Circuits and Systems (ISCAS)* 2010:253–256
 23. Chen H, Hwang BJ, Shyr J, Liu T (2020) The effect of different deep network architectures upon CNN-based gaze tracking. *Algorithms* 2020:1–18
 24. Zhang G-W, Jien Kato Yu, Wang KM (2014) Multi-stage deep convolutional learning for people re-identification. *Computer System Science & Engineering* 2014:265–274
 25. Zheng Qiumei; Tan Dan; Wang Fenghua (2020) Improved Convolutional Neural Network Based on Fast Exponentially Linear Unit Activation Function. *IEEE Access* 2020:151359–151367
 26. Chen J, Hua Z, Wang J, Cheng S (2017) A convolutional neural network with dynamic correlation pooling. *International Conference on Computational Intelligence and Security* 2017:496–499
 27. Vatalaro M, Moposita T, Strangio S, Trojman L, Vladimirescu A, Lanuzza M, Crupi F (2021) A low-voltage, low-power reconfigurable current-mode softmax circuit for analog neural networks. *Electronics* 2021:1–11
 28. Armato SG, McLennan G, Bidaut L (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* 38:915–931
 29. Krewer H, Geiger B, Hall L. O (2013) Effect of texture features in computer aided diagnosis of pulmonary nodules in low dose computed tomography, *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2013:3887–3891.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



R. Manickavasagam received a B.E. degree in Electronics and Communication Engineering from the Manonmaniam Sundaranar University, Tirunelveli, India, in 2000 and an M.E. degree in Applied Electronics from the Sathyabama University, Chennai, India, in 2008. He received a Ph.D. degree in Electronics and Communication Engineering from Anna University, Chennai, India, in 2021. He has 20 years of teaching

experience. He is currently an Associate Professor in the Department of Biomedical Engineering at Alpha College of Engineering, Chennai, India. He has published more than 30 papers in international and national journals and conference proceedings. His areas of research include medical image processing, deep learning, soft computing, and computer networks.



Srinivasan Selvan received a B.E. degree in Electronics and Communication Engineering and an M.E. degree in Communication Systems from the University of Madras, Chennai, India, in 1977 and 1979, respectively, and a Ph.D. degree in Computer Science and Engineering from the Madurai Kamaraj University, Madurai, India, in 2001. He has 41 years of teaching experience. He is currently a Professor in the Department of Electronics and Communication Engineering at St. Peter's Institute of Higher

Education and Research, Chennai, India. He has published more than 300 papers in international and national journals and conference proceedings. His areas of research include digital image processing, deep learning, soft computing, digital signal processing, and computer networks.



Mary Selvan received a B.E. degree in Electrical and Electronics Engineering from the Bharathiar University, Coimbatore, India, in 2004 and an M.E. degree in Computer Science and Engineering from Anna University, Chennai, India, in 2009. She is currently pursuing a Ph.D. degree from Anna University, Chennai, India. She has 16 years of teaching experience. She is currently an Associate Professor in the Department of Computer Science and Engineering, St. Peter's College of Engineering and Technology, Chennai, India. She has published more than 11 papers in international and national journals and conference proceedings. Her areas of research include digital image processing, deep learning, and computer networks.

neering and Technology, Chennai, India. She has published more than 11 papers in international and national journals and conference proceedings. Her areas of research include digital image processing, deep learning, and computer networks.