Journal of Engineering Science and Technology
Vol. 16, No. 4 (2021) 3250 – 3266

CLASSIFICATION OF BENIGN AND MALIGNANT PULMONARY NODULES IN LDCT IMAGES USING RADIOMIC FEATURES

SHABANA R. ZIYAD[1,*], V. RADHA[2], THAVAVEL VAYYAPURI[1]

[1]Department of Computer Science, College of Computer Engineering and Sciences, Prince Sattam Bin Abdulaziz University, Al Kharj, Kingdom of Saudi Arabia
[2]Department of Computer Science, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, India
[1]Department of Computer Science, College of Computer Engineering and Sciences, Prince Sattam Bin Abdulaziz University, Al Kharj, Kingdom of Saudi Arabia
*Corresponding Author: ziyadshabana@gmail.com

Abstract

Cancer is a cause of premature death in humans. Cancer-related deaths are increasing worldwide over the years. Among all cancers, lung cancer contributes to a major proportion of deaths. The survival rate of patients with lung cancer can, however, be increased to a great extent by timely and efficient diagnosis and treatment. The present study aimed to develop a computer-aided diagnosis system with a novel segmentation method and adoption of radiomic features to classify benign and malignant nodules. Initial pre-processing was performed by wavelets for denoising the image. The novel segmentation method based on fuzzy c-means clustering enhanced by genetic algorithm segments of the lung region in the low-dose computed tomography images. The population size was initialized with cluster centres randomly. The fitness value was calculated by minimizing the individual pixel distance from the cluster centres. The best parents were selected according to the fitness value. The process of crossover and mutation was applied to the parent chromosomes. These steps were repeated for a finite number of times to obtain the optimum solution. Nodules were extracted by morphological operations of retrieving the objects based on their size. Radiomic features were extracted from the images, and these feature sets were reduced to 12 features by applying the least absolute shrinkage and selection operator (LASSO) for feature selection. Benign and malignant nodules in the low dose computed tomography (LDCT) images were classified using the ensemble method of AdaBoost. The sensitivity, specificity, and accuracy achieved by this method were 91.66%, 100%, and 98%, respectively.

Keywords: Computer-aided diagnosis, Features, Fuzzy c-means clustering, LASSO, Lung cancer, Radiomic.

1. Introduction

Lung cancer originates from the cells lining the air passages and can be broadly classified as small cell lung cancer and non-small cell lung cancer. Non-small cell lung cancer accounts for 80% of the lung cancer cases. Lung cancer is the second most common cancer among men in India. The number of lung cancer cases is much higher in males than in females. The reason is the higher use of tobacco in men than in women. The likelihood of a smoker developing lung cancer as compared to that for a non-smoker is 20:1 [1].

Common diagnostic tests for lung cancer include imaging tests, sputum cytology, and tissue sample testing or biopsy. Low-dose computed tomography (LDCT) imaging test is a commonly suggested imaging test for patients at a high risk of developing lung cancer. The biopsy can be performed as bronchoscopy or mediastinoscopy [2]. In both these procedures, the patient experiences pain and discomfort. Computer-aided diagnosis (CADx) tools have emerged as an aid to radiologists in diagnosing lung nodules and classifying them as benign or malignant. Consequently, extensive research studies have been conducted in the area of CADx. There is much scope for improvement in the phase of classification.

According to previous literature, the performance of a classifier solely depends on the quality and quantity of extracted features. Radiomics is a method of extracting a large amount of mineable medical features from clinical images. Radiomic features can reveal many characteristics that can accurately classify the nodules as benign or malignant. The CADx used in the present study performs pre-processing of LDCT images by using wavelets with soft thresholding to remove noise and improve contrast. This step is followed by segmentation of the lung by a novel fuzzy c-means clustering enhanced by a genetic algorithm (GAFCM). The segmented images then undergo morphological operations to segment the lung nodules. The radiomic features for these lung nodules are then extracted. Feature selection followed by the ensemble method reduces the set of radiomic features and subsequently improves the efficiency of the CADx.

The present study aimed to develop CADx based on radiomic features which uses the feature selection method to obtain the best discriminating features.

2. Related Work

Extensive studies have been conducted by several researchers in the field of developing CADe and CADx for the detection of lung nodules. An improved classification performed by stacked denoising auto-encoder (SDAE)-based CADx algorithm has been reported [3]. Three sets of features are extracted from Gabor filter, local binary pattern (LBP), and texture features. Signed distance along with LBP generates a combination of shape and texture features which adopt the ensemble method of support vector machine (SVM) and K-nearest neighbour (KNN) for classifying nodules in chest computed tomography [4].

The lung nodules are classified as benign or malignant through an adaptive weighting scheme learned with error back propagation that uses the heterogeneity of voxel values and shape of the nodules [5]. A CADx in which SVM was adopted for classifying lung nodules considered the shape and texture features for classification [6]. In another study, five classifiers, namely SVM, random forest (RF), linear regression (LR), extreme learning machine (ELM), and KNN, were

evaluated for their performance on classifying lesions extracted from the CT images of patients with lung cancer. Initially, the feature selection method was performed using the relief feature selection algorithm. Subsequently, the prediction results of the five classifiers were combined to predict the malignancy of the lesions in CT images [7, 8].

A CADx scheme was developed to classify nodules and to detect the advanced stage of malignant nodules in the lung region by feature extraction. Approximately sixty-six 3D features were selected, and the classifiers SVM, naive Bayes, and linear discriminant analysis (LDA) were trained separately using three types of datasets. Shape, texture, and intensity features were adopted to classify the lung nodules. The classification was performed by a two-layer artificial neural network [9]. Safta et al. [10] proposed the concept of multiple instances learning where classification was performed using SVM.

Radiomic features were adopted to classify cancer as adenocarcinoma and squamous cell carcinoma [11]. Lee et al. [12] used the features of tumour size and shape for feature extraction. Features showing high reproducibility in voxel settings and binning settings were detected. The least absolute shrinkage and selection operator (LASSO) was used for feature reduction. The RF classifier validated the efficiency of this method.

The classification of benign and malignant lung nodules based on the statistical and texture features was best achieved with the multilayer perceptron classifier with an accuracy of 88.55% [13]. A novel ensemble shape gradient features (NESGF) descriptor was proposed for lung nodule classification. The classification was performed using the histogram of oriented surface normal vectors (HOSNV) and multi-coordinate histogram of gradient descriptor. The RF method has been adopted to classify nodules by applying the ensemble concept [14]. A novel method was proposed by authors where the 3D labelling technology is adopted to classify the nodules into two groups. A rule-based classifier is used to classify true nodules for the non-vessel group. Dot filters are used for the vessel tree group.

The novel feature selection approach selects the discriminating features. The weighted SVM (WSVM) is adopted to classify vessels and nodules [15]. The lung features were analysed by discrete AdaBoost-optimized ensemble learning-generalized neural networks for public lung database images, which classified the features with more precision [16]. Contrast enhancement was performed by the gamma correction max intensity weights approach. The multiple texture, point, and geometric features were extracted from enhanced images, which were applied to a serial canonical correlation-based fusion.

Feature selection by weighted neighbourhood component analysis (NCA) followed by ensemble classifier was adopted for final classification. The method achieved sensitivity, specificity, and accuracy of 86%, 90%, and 94%, respectively [17]. Saba [18] proposed a detection and classification method based on the voting of classifiers.

3. Materials and Methods

The CADx in this study has six phases: pre-processing, segmentation, nodule detection, nodule extraction, feature extraction, and classification. The lung LDCT images are extracted from the Lung Image Database Consortium-Image Database

Resource Initiative (LIDC-IDRI) database. The proposed methodology of developing CADx in this study is shown in Fig. 1. In this study, 600 LDCT images were acquired for the classification of nodules as benign or malignant. Figure 2 shows the flowchart for the proposed CADx based on the novel segmentation technique, nodule extraction, feature selection, and classification.

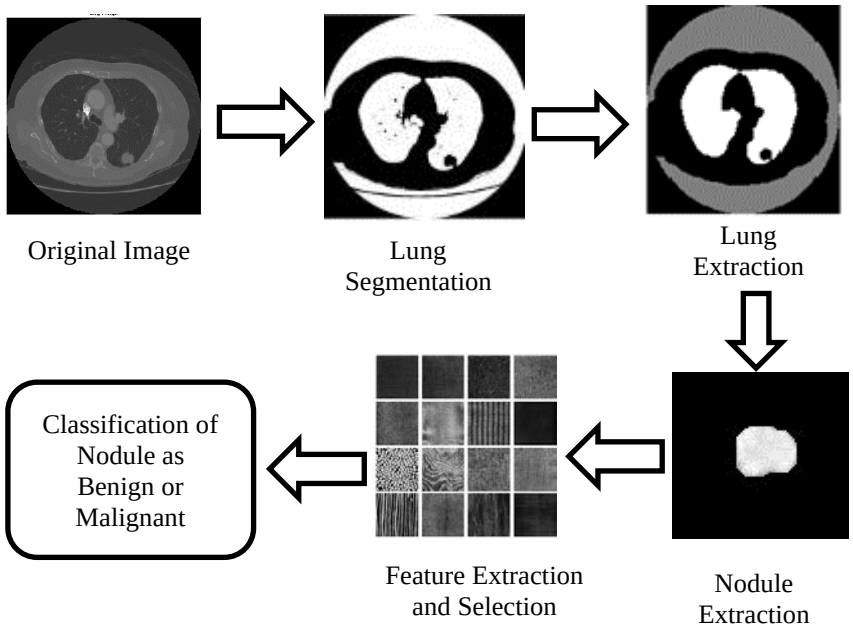

Fig. 1. Phases of the proposed CADx for lung nodule classification.

3.1. Dataset

The CT images are retrieved from the LIDC-IDRI public database. The dataset consists of 600 CT images from the LDCT-IDRI public lung database, with 95 images showing malignant nodules [19]. Among the 600 LDCT images, a set of 550 LDCT images form the training set and a set of 50 LDCT images form the testing set. The ratio of training set to testing set is 11:1.

3.2. Pre-processing

In the pre-processing phase, the images undergo a denoising process with wavelets. The wavelets can be localized in time and space. Haar wavelets are applied for denoising with soft thresholding. This threshold value is dependent on the image size and standard deviation of noise. The universal threshold used is given in Eq. (1).

$$T = \sigma\sqrt{2\log N \times N} \tag{1}$$

where N is the image size and $\sigma = \frac{Median\left|W_1^j\right|}{0.6745}\ \forall\ i\ diagonal\ coefficients$ where W_1^j is the wavelet coefficient in scale 1 [20]. $T' = (W_1^j)$. After applying the

threshold inverse, discrete wavelet transform (DWT) is applied to retrieve the denoised image.

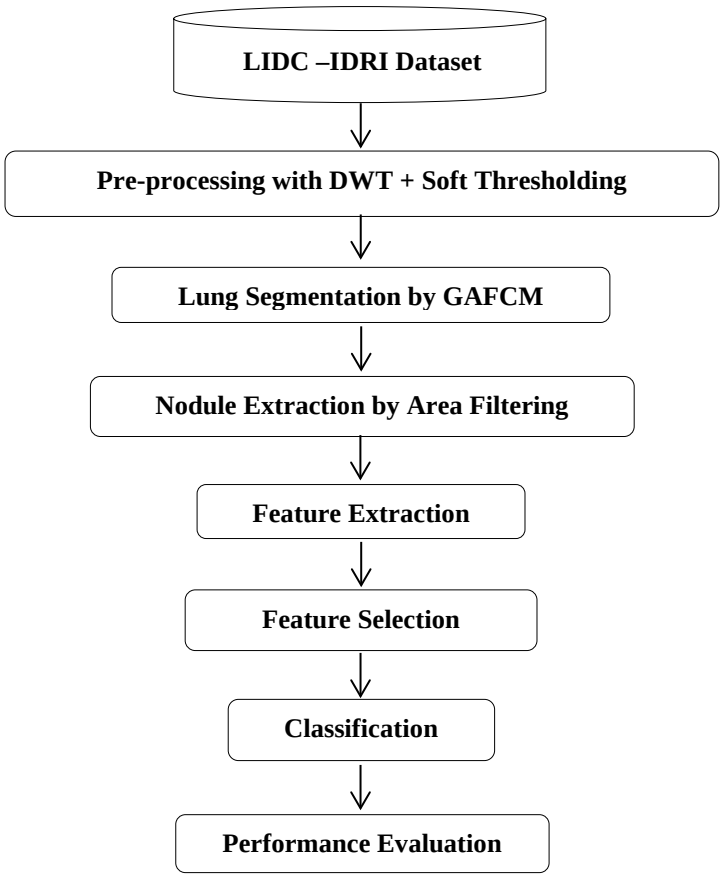

Fig. 2. Flow chart of the proposed method.

3.3. Segmentation

In medical imaging, segmentation is the process of retrieving the region of interest in any image. In this study, segmentation refers to the extraction of the lung parenchyma region to classify the lung nodules as benign or malignant. Lung segmentation is performed by a novel method in which the fuzzy c-means clustering technique is enhanced by genetic algorithm (GAFCM). Fuzzy c-means clustering is a method of clustering that allows the pixels to group into one or more clusters. It aims to minimize the objective function given in Eq. (2).

$$TWCSE(\delta) = \sum_{i=1}^{N} \sum_{j=1}^{c} u_{ij}^{m} \left\| x_i - c_j \right\|^2 \quad (2)$$

where

$$c_j = \frac{\sum_{j=1}^{N} u_{ij}^{m} . x_i}{\sum_{j=1}^{N} u_{ij}^{m}} \quad (3)$$

$$u_{ij} = \frac{1}{\sum_{k=1}^{c}\left(\frac{\|x_i - c_j\|}{\|x_i - c_k\|}\right)^{\frac{2}{m-1}}} \quad (4)$$

where u_{ij} is the degree of membership of $x_i$in cluster j. x_i is the data point, c_j is the cluster centre, and $\|x_i - c_j\|$ is the distance between the data point and the cluster centre. The steps are repeated until the cluster centre converges. A genetic algorithm (GA) is a heuristic algorithm based on Charles Darwin's theory of natural evolution. The possible cluster centres are generated randomly and form the population that comprises a set of chromosomes. Each possible cluster centre is termed as a chromosome. A fitness function is defined that evaluates the fitness of the chromosome to be a parent for the new offspring. The fitness function optimizes the inter-cluster distance between the cluster centre and the pixels of each cluster. The fittest cluster centre is selected, and it undergoes the process of crossover and mutation. The process is repeated until the cluster centre converges.

This novel method improves the accuracy of lung segmentation, particularly in identifying the juxtapleural nodules. This segmentation process is followed by the morphological closing operation. This morphological closing is essential to identify the juxtapleural nodules attached to the lung boundaries.

3.4. Nodule extraction

Nodule extraction is a crucial phase in the lung CADx. This phase demarcates the lung nodules from similar-looking structures such as blood vessels. The reason for the high false-positive rate is similar appearance of the nodules to the blood vessels present in the pulmonary region. Hence, the lung nodule extraction phase aims to reduce false positives in detecting lung nodules for any CADx.

The morphological operation is the most widely used methodology in nodule extraction in previous studies as it is known to preserve the juxtapleural nodules. Hence, the same methodology was used in the present study. All the nodules are extracted by the edge detection method followed by morphological operations. The image is first converted into a binary image. The merged components in the image are demarcated, and the nodules with an area of the specified range are then extracted. As the radius of the pulmonary nodule is usually in the range of 1 to 3 mm, the area can be calculated to extract the nodules from the LDCT image. The boundaries are also plotted for the extracted nodules. The extracted nodule area is marked on the LDCT Dicom Image. The algorithm for the proposed CADx for lung nodule classification is shown in Fig. 3.

3.5. Feature extraction

The features used in this study are histogram features, morphological features, texture features, and ROI-based features. Histogram features include root mean square (RMS), mean, median, standard deviation, variance, skewness, kurtosis, energy, entropy, and mode. Morphological features include real size, convex area, area, perimeter, compactness, eccentricity, solidity, orientation, and equivdiameter.

Texture features are statistically measured features. Gray-level co-occurrence matrix (GLCM) is a statistically measured feature based on the gray length co-occurrence matrix which specifies the frequency of occurrence of the neighbouring pixels with homogenous intensity.

GLCM features include contrast, correlation, energy, homogeneity, entropy, autocorrelation, cluster prominence, cluster shade, dissimilarity, sum of squares, sum average, sum variance, difference variance, difference entropy, maximal correlation coefficient, inverse difference (ID), inverse difference normalized (IDN), and inverse difference moment normalized features (IDMN).

Gray-level run length matrix (GLRLM) includes 11 features as shown in Table 1. ROI features are the region of interest-based features. Irrelevant and redundant features affect the performance of the classification model. Hence, the performance and accuracy of the model are improved by reducing the feature set with the feature selection techniques.

Step 1: Select dataset.
Step 2: Apply DWT and soft thresholding on the LDCT image to obtain approximation and detail sub-bands.
Step 3: Define cluster centre for each image.
Step 4: Initialize the GA parameter population size, crossover probability, and termination criterion.
Step 5: Generate a population with random cluster centres.
Step 6: While (TWCSE)
for i = 1 to PopSize do
for each individual
Calculate the fitness value using

$$Objective\ function(\delta) = \min\ (TWCSE(\delta))$$

where

$$TWCSE(\delta) = \sum_{k=1}^{K} \sum_{x_i=C_k}^{n} \|X_i - Z_k\|^2$$

$$Z_k = \frac{\sum_{x_i \in C_k} X_i}{|C_k|}$$

$|C_k|$ denotes the number of pixels in C_k, $(Z_1, Z_2, \ldots, Z_K)$ denote the target cluster centres.

Sort the population according to the ranking of fitness function.
Apply crossover, create offspring, and mutate offspring.
End of for loop.
Step 7: Replace the old population with the new ones.
Step 8: End of while loop.
Step 9: Cluster the feature space using the GA-optimized centres.
Step 10: Perform lung segmentation with optimized centres.
Step 11: Extract nodules from the segmented lung region by morphological operations.

Fig. 3. Algorithm for the proposed CADx.

Table 1. Summary of features.

Features Set	No. of Features	Features
Histogram	10	RMS, mean, median, standard deviation, variance, skewness, kurtosis, energy, entropy, mode
Morphological Features	9	Real size, convex area, area, perimeter, compactness, eccentricity, solidity, orientation, equivdiameter
GLCM	22	Contrast, correlation, autocorrelation, cluster prominence, cluster shade, dissimilarity, energy, entropy, homogeneity, maximum probability, sum of squares-variance, sum average, sum variance, sum entropy, difference variance, difference entropy, information measure of correlation, maximal correlation coefficient, inverse variance, inverse difference normalized, information measure of correlation, inverse difference moment normalized
GLRLM	11	Long run emphasis Gray-level non-uniformity Run length non-uniformity Run percentage Low gray-level run emphasis High gray-level run emphasis Short run low gray-level emphasis Short run high gray-level emphasis Long run low gray-level emphasis Long run high gray-level emphasis
ROI-based Features	7	Mean of intensity value, SD of intensity value, mean of covariance, SD of covariance, mean of inside region, SD of annular region, roughness

3.6. Feature selection

The efficiency of machine learning algorithms is improved by providing distinct and relevant input features. Redundant and irrelevant features can decrease the performance of the classification model. Therefore, the process of feature reduction not only reduces the training time, but it also reduces the evaluation time as well as complexity and improves accuracy. Feature selection methods are broadly classified as Filter Methods, Wrapper Methods, and Embedded Methods. In filter-based methods, the features which rank below a certain threshold are removed from the selected features of the subset. Filter methods are dependent on statistical tests such as chi-square and analysis of variance (ANOVA) for determining the correlation coefficients of different features with the outcome variable.

In the wrapper method, evolutionary algorithms generate the different subsets of features, which are then evaluated by the performance of the predictors selected using a supervised learning algorithm. Wrapper methods include forward selection, backward selection, and recursive feature elimination. Forward selection starts with the empty subset of features and iteratively adds relevant features. Backward

selection initially adds all the features and then removes any irrelevant features in each of the iterations. The recursive feature elimination method eliminates the features based on the greedy optimization technique. The embedded method is a combination of filter and wrapper methods.

In the present study, we use the LASSO method which is an embedded method for feature reduction. LASSO applies a shrinking or regularization process where it evaluates the coefficients of the variables, shrinking many of them to zero. The variables that have non-zero coefficients after the above process are selected as part of the model. Variables that are strongly related to response variables have non-zero coefficient values. Features whose coefficients reduce to zero are removed from the feature set. This prevents the model from yielding biased results [21]. The features extracted include morphological features of real size, area, perimeter, and solidity; GLCM features of energy, cluster shade, autocorrelation, entropy, and homogeneity; GLRLM features of run-length non-uniformity and long-run high gray-level emphasis; and ROI-based feature of roughness.

3.7. Classification

Recently, ensemble learning-based classification methods have been used in the field of medical imaging. In ensemble methods, the classifier training process has feature selection as one of its components. SVM and decision tree are the popularly used classifiers in this ensemble method [22]. The ensemble classification model generates a set of classification results by using multiple weak classifiers. These multiple prediction or classification results are mapped to the final results by using the voting schemes.

The Bagging method generates sample subsets by randomly sampling from the training dataset and then uses the obtained subsets to train the basic models for integration. The training of basic models in the Bagging method is performed in a parallel manner [23]. If the dataset has n samples, then each weak learner is given a random sample of the dataset as input. The weak learners take the input dataset which can be named as $D^1, D^2, D^3 \ldots D^M$ where M is the number of weak learners. These samples $D^1, D^2, D^3 \ldots D^M$ for the weak learners are generated by the random sampling method with replacement from the original dataset. This indicates that some samples of the dataset may be repeated in the training data chosen for another weak learner. The probability of element appearing in every dataset is the same. This step is called bootstrap, and the subsequent step of consolidating the output of the weak learners is known as aggregation. The outcome for prediction problems is the average of the outcome predictive values of the weak learners. The result of classification is computed by the majority voting scheme.

In the RF technique, predictors for the decision split are selected randomly by trees in ensemble. In this method, the number of predictors selected randomly for each split is equal to the square root of the number of predictors for classification. As a method to improve the prediction of binary classification and to develop a strong classifier, AdaBoost was introduced by Freund and Schapire [24].

AdaBoost primarily aims to train weak classifiers and combine them to form a final classifier. The weight of each sample is calculated depending on its ease of classification, and the data distribution is altered accordingly to improve the accuracy of overall classification. In AdaBoost, the weak classifiers connected in

series improve the classification accuracy by learning from the samples that are misclassified in the previous classifier. This decision tree classifier is known as stumps. At any stage, an individual tree is trained to learn the weakness of the previous tree. The misclassified sample is given more weightage and is boosted such that the subsequent trees focus on accurately classifying the misclassified sample. It has been proven that the addition of more classifiers in the series improves the classification accuracy by giving more weightage to the misclassified samples [25]. In the present study, we compared three classifiers, namely SVM, Bagging, and AdaBoost.

4. Performance Metrics

The traditional method of SVM and ensemble methods of bagging and AdaBoost are trained with a dataset of 550 LDCT images, with 95 images consisting of malignant nodules. Twenty-three images have nodules greater than 3 mm, and the remaining images have nodules of size less than 3 mm. The performance metrics for the classification methods include specificity, sensitivity, accuracy, false-positive rate, and receiver operating characteristic (ROC). The formula for specificity, sensitivity, and accuracy is given in Eqs. (5) – (7) [26]. TN is True Negative, TP is True Positive, FN is False Negative, and FP is False Positive.

$$\text{Specificity} = TN/(TN+FP) \tag{5}$$

$$\text{Sensitivity} = TP/(TP+FN) \tag{6}$$

$$\text{Accuracy} = (TP+TN)/(TP+TN+FP+FN) \tag{7}$$

5. Results and Discussion

All image processing algorithms were implemented with MATLAB 2018b. In the present study, after the pre-processing of the images, the lung region was segmented with a novel segmentation method, and the nodules were extracted using the morphological operation. In the feature extraction phase, a total of 59 features of the lung nodules were considered for the initial study. Because a large number of features can affect the performance of a classification model, feature reduction was performed by identifying the correlation among and the discriminating power of the different features.

In the present study, a t-test feature set was initially used under the hypothesis that 'there is no interaction between the features', and a graph was derived for the p-value as shown in Fig. 4. The figure clearly shows that up to 80% of the features have a *p*-value of less than 0.05, and these features have a strong discriminating power. Table 2 shows the results of sensitivity, specificity, and accuracy achieved for the three ensemble classification methods of SVM, bagging, and AdaBoost for the complete feature set. The reduced sensitivity of the above methods is due to the increased false negatives occurring in the results of the classification model. The LASSO method reduces the feature set to 12 features.

Table 3 presents the results of sensitivity, specificity, and accuracy achieved for the three ensemble classification methods of SVM, bagging, and AdaBoost with the reduced feature set. The ROC was constructed for the methods of SVM, Bagging, and AdaBoost for the complete feature set shown in Figs. 5, 6, and 7, respectively. The receiver operating characteristic curve was constructed for the

methods of SVM, Bagging, and AdaBoost after feature reduction as shown in Figs. 8, 9, and 10, respectively. The reduced feature set when considered as an input for classification using the traditional SVM method achieved sensitivity, specificity, and accuracy of 83.33%, 100%, and 96%, respectively. The specificity, sensitivity, and accuracy for using the same feature subset were 100%, 91.66%, and 98%, respectively, for the Bagging method and 100%, 91.66%, and 98%, respectively, for the AdaBoost method. Figures 11 and 12 show the confusion matrix for the two ensemble methods of Bagging and AdaBoost for the reduced feature set.

Table 2. Comparison of the SVM, Bagging, and AdaBoost methods for the classification of lung nodules in LDCT images using the full set of features.

Method	Sensitivity (%)	Specificity (%)	Accuracy (%)
SVM	50	100	90
Bagging	54	100	90
AdaBoost	63.63	100	92

Table 3. Comparison of the SVM, Bagging, and AdaBoost methods for the classification of lung nodules in LDCT images for the reduced feature set.

Method	Sensitivity (%)	Specificity (%)	Accuracy (%)
SVM	83.33	100	96
Bagging	91.66	100	98
AdaBoost	91.66	100	98

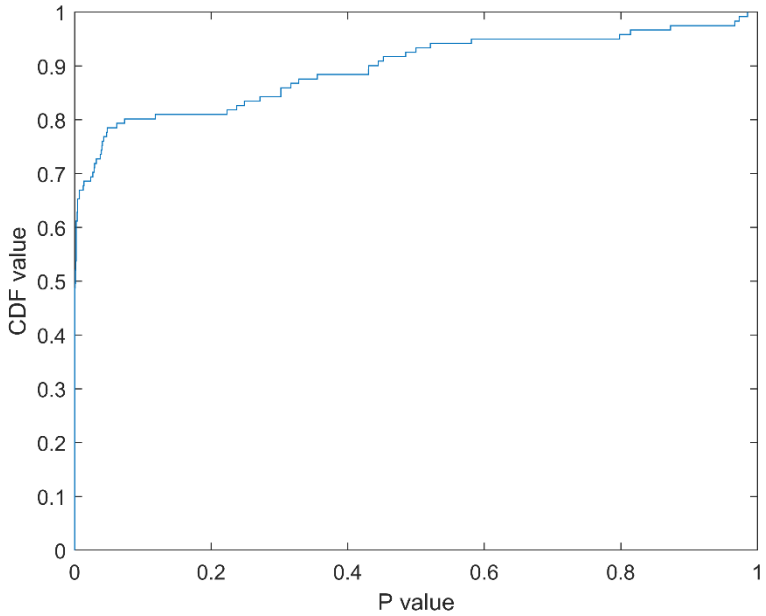

Fig. 4. P-value graph for the feature set.

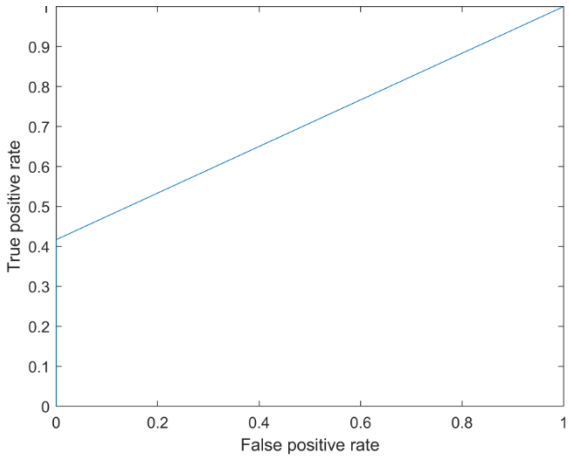

Fig. 5. ROC for SVM classifier for all features.

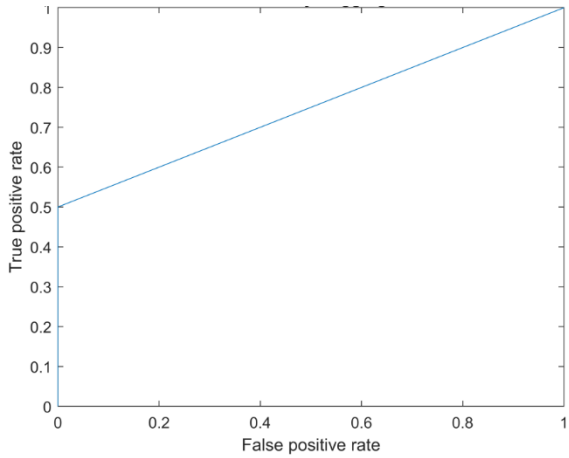

Fig. 6. ROC for Bagging for all features.

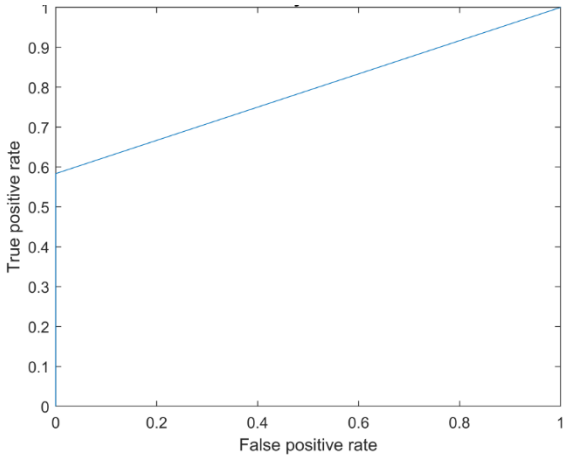

Fig. 7. ROC for AdaBoost for all features.

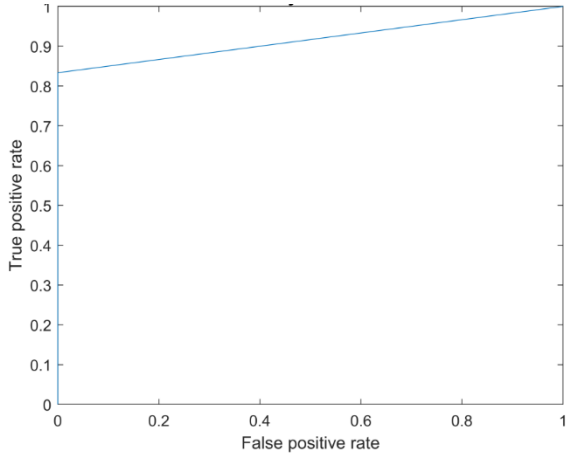

Fig. 8. ROC for SVM for reduced feature subset.

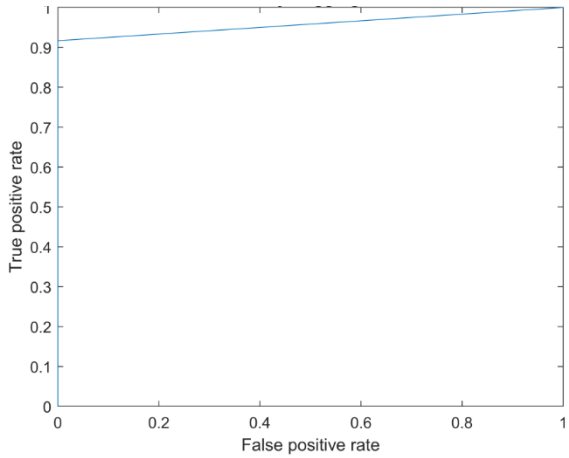

Fig. 9. ROC for the Bagging ensemble method for the reduced feature subset.

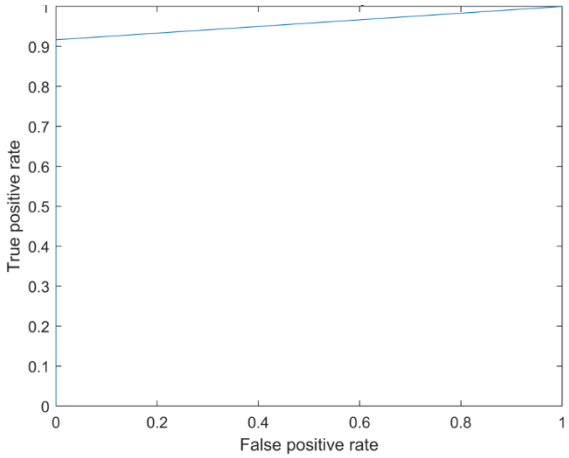

Fig.10. ROC for the AdaBoost ensemble method for the reduced feature subset.

TP=10	**FP=0**
FN=2	**TN=38**

Fig. 11. Confusion matrix for the SVM method for the classification of lung nodules in LDCT images for the reduced feature set.

TP=11	**FP=0**
FN=1	**TN=38**

Fig. 12. Confusion matrix for the Bagging and AdaBoost methods for the classification of lung nodules in LDCT images for the reduced feature set.

6. Conclusions

The present study identified the feature subset with a high discriminating power to classify lung nodules as benign or malignant. The false-positive rate is considerably reduced by the reduction of the feature set using the LASSO feature reduction method. The novel method of lung segmentation with the combination of feature reduction and classification by the ensemble method yielded much higher performance than that achieved using other state-of-art methods.

Acknowledgements

We would like to express our sincere gratitude to Dr. Haya Alaskar, Vice Dean, College of Computer Engineering and Sciences (Female Section), Prince Sattam Bin Abdulaziz University, Al-Kharj, KSA, for her constant support and valuable guidance in the research work. This publication was supported by Deanship of Scientific Research at Prince Sattam Bin Abdulaziz University, Al-Kharj, KSA.

Nomenclatures

C_j	Cluster centre
M	Number of weak learners
N	Image size
P	Statistical measure
T	Universal threshold
u_{ij}	Degree of membership
W	Wavelet coefficient
X	Set of data points
Z	Set of pixels in a cluster

Greek Symbol

σ	Average variance of noise

Abbreviations

ANOVA	Analysis of variance
CADe	Computer-aided detection
CADx	Computer-aided diagnosis
ELM	Extreme learning machine
FN	False negative

FP	False positive
GA	Genetic algorithm
GAFCM	Fuzzy c-means clustering by genetic algorithm
GLCM	Gray-level co-occurrence matrix
GLRLM	Gray-level run length matrix
HOSNV	Histogram of oriented surface normal vectors
ID	Inverse difference
IDMN	Inverse difference moment normalized features
IDN	Inverse difference normalized
KNN	K-nearest neighbour
LASSO	Least absolute shrinkage and selection operator
LBP	Local binary pattern
LDA	Linear discriminant analysis
LIDC-IDRI	Lung Image Database Consortium-Image Database Resource Initiative
LDCT	Low-dose computed tomography
LR	Linear regression
NCA	Neighbourhood component analysis
NESGF	Novel ensemble shape gradient features
RF	Random forest
RMS	Root mean square
ROC	Receiver operating characteristic
SDAE	Stacked denoising auto-encoder
SVM	Support vector machine
TN	True negative
TP	True positive
WSVM	Weighted support vector machine

References

1. India Against Cancer. (n.d.). Lung cancer. Retrieved May 10, 2020, from http://cancerindia.org.in/lung-cancer/.
2. London Health Sciences Centre. (n.d.). Thoracic surgery. Retrieved May 10, 2020, from https://www.lhsc.on.ca/thoracic-surgery/thoracic-surgery-83
3. Cheng, Z.-H.; Ni, D.; Chou, Y.-H.; Qin, J.; Tiu, C.-M.; Chang, Y.-C.; Huang, C.-S.; Shen, D.; and Chen, C.-M. (2016). Computer-aided diagnosis with deep learning architecture: applications to breast lesions in US images and pulmonary nodules in CT scans. *Scientific Reports*, 6, 24454.
4. Farag, A.A.; Ali, A.; Elshazly, S.; and Farag. A.A. (2017). Feature fusion for lung nodule classification. *International Journal of Computer Aided Radiology and Surgery*, 12(10), 1809-1818.
5. Xie, Y.; Xia, Y.; Zhang, J.; Feng, D.D.; Fulham, M.; and Cai, W. (2017). Transferable multi-model ensemble for benign-malignant lung nodule classification on chest CT. In: Descoteaux M., Maier-Hein L., Franz A., Jannin P., Collins D., Duchesne S. (Eds.). *International Conference on Medical Image Computing and Computer Assisted Intervention − MICCAI 2017*. Lecture Notes in Computer Science, Vol. 10435. Springer, Cham.

6. Shaukat, F.; Raja, G.; Gooya, A.; and Frangi, A.F. (2017). Fully automatic and accurate detection of lung nodules in CT images using a hybrid feature set. *Medical Physics*, 44(7), 3615-3629.
7. Shen, Y.; Xu, F.; Zhu, W.; Hu, H.; Chen, T.; and Li, Q. (2020). Multiclassifier fusion based on radiomics features for the prediction of benign and malignant primary pulmonary solid nodules. *Annals of Translational Medicine*, 8(5), 171.
8. Gong, J.; Liu, J.Y.; Sun, X.W.; Zheng, B.; and Nie, S.D. (2018). Computer-aided diagnosis of lung cancer: the effect of training data sets on classification accuracy of lung nodules. *Physics in Medicine & Biology*, 63(3), 035036.
9. Al-Shabi, M.; Lan, B.L.; Chan, W.Y.; Ng, K.; and Tan, M. (2019). Lung nodule classification using deep local–global networks. *International Journal of Computer Assisted Radiology and Surgery*, 14(10), 1815-1819.
10. Safta, W.; Farhangi, M.M.; Veasey, B.; Amini, A.; and Frigui, H. (2019). Multiple instance learning for malignant vs. benign classification of lung nodules in thoracic screening CT data. *Proceedings of 16TH IEEE International Symposium on Biomedical Imaging*, Venice, Italy, 1220-1224.
11. Lu, L.; Li, L.; Yang, H.; Schwartz, L.H.; and Zhao, B. (2019). Radiomics for classifying histological subtypes of lung cancer based on multiphasic contrast-enhanced computed tomography. *Journal of Computer Assisted Tomography*, 43(2), 300-306.
12. Lee, S.-H.; Cho, H.-H.; Yun, L.H.; and Park, H. (2019). Clinical impact of variability on CT radiomics and suggestions for suitable feature selection: a focus on lung cancer. *Cancer Imaging*, 19(1), 54.
13. Singh, G.A.P.; and Gupta, P.K. (2019). Performance analysis of various machine learning-based approaches for detection and classification of lung cancer in humans. *Neural Computing and Applications*, 31(10), 6863-6877.
14. Jaffar, M.A.; Zia, M.S.; Sultan, M.; Hussain, M.; Siddiqui, A.B.; and Jamil, U. (2020). An ensemble shape gradient features descriptor based nodule detection paradigm: a novel model to augment complex diagnostic decisions assistance. *Multimedia Tools and Applications*, 79(13), 8649-8675.
15. Gu, Y.; Lu, X.; Zhang, Z.; Zhao, Y.; Yu, D.; Gao, L.; Cui, G.; Wu, L.; and Zhou, T. (2019). Automatic lung nodule detection using multi-scale dot nodule-enhancement filter and weighted support vector machines in chest computed tomography. *PLoS One*, 14(1), e0210551.
16. Shakeel, P.M.; Tolba, A.; Al-Makhadmeh, Z.; and Jaber, M.M. (2020). Automatic detection of lung cancer from biomedical data set using discrete AdaBoost optimized ensemble learning generalized neural networks. *Neural Computing & Application*, 32(3), 777-790.
17. Attique, M.; Khana, S.; Asifa Kashif, R.; Sharif, M.I.; Muhammad, N.; Shahd, J.H.; Zhang, Y.D.; and Satapathy, S.D. (2020). Lungs cancer classification from CT images: an integrated design of contrast based classical features fusion and selection. *Pattern Recognition Letters*, 129, 77-85.
18. Saba, T. (2019). Automated lung nodule detection and classification based on multiple classifiers voting. *Microscopy Research and Technique*, 82(9), 1601-1609.
19. Vendt, B. (n.d.). LIDC-IDRI. Retrieved May 14, 2020, from https://wiki.cancerimagingarchive.net/display/Public/LIDC-IDRI.

20. Naveed, K.; Shaukat, B.; Ehsan, S.; Mcdonald-Maier K.D.; and ur Rehman, N. (2019). Multiscale image denoising using goodness-of-fit test based on EDF statistics. *PLoS One*, 14(5), e0216197.
21. Rawale, S. (2018). Feature selection methods in machine learning. Retrieved May 14, 2020, from https://medium.com/@sagar.rawale3/feature-selection-methods-in-machine-learning-eaeef12019cc.
22. Chen, C.W.; Tsai, Y.H.; Chang, F.R.; and Lin, W.C. (2020) Ensemble feature selection in medical datasets: Combining filter, wrapper, and embedded feature selection results. *Expert Systems*, 37(5), e12553.
23. Dong, X.; Zhiwen Y.U.; Wenming C.A.O.; Yifan S.H.I.; and Qianli M.A. (2020). A survey on ensemble learning. *Frontiers of Computer Science*, 14(2), 241-258.
24. Freund, Y.; and Schapire, R.E. (1996). Game theory, on-line prediction and boosting. *Proceedings of the Ninth Annual Conference on Computational Learning Theory*, Barcelona, Spain, 325-332.
25. Machova, K.; Puszta, M.; Barcák, F.; and Bednár, P. (2006). A comparison of the bagging and the boosting methods using the decision trees classifiers. *Computer Science and Information Systems*, 3(2), 57-72.
26. Baratloo, A.; Hosseini, M.; Negida, A.; and El Ashal, G. (2015). Part 1: Simple definition and calculation of accuracy, sensitivity and specificity. *Emerg (Tehran)*, 3(2), 48-49.