ASEAN Engineering Journal

Full Paper

ROC CURVE ANALYSIS OF DIFFERENT HYBRID FEATURE DESCRIPTORS USING MULTI CLASSIFIERS

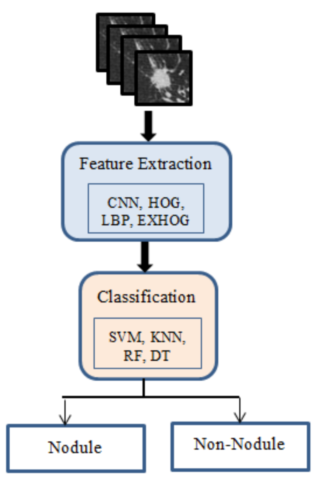
S. Piramu Kailasam[a]*, E. Siva Sankari[b], R. Kumuthini[c]

[a]Department of Computer Science, Sadakathullah Appa College (Affiliated to Manonmaniam Sundaranar University, Tirunelveli), Palayamkottai Tirunelveli, Tamilnadu, India
[b] Department of Computer Science and Engineering, Government College of Engineering, Tirunelveli, Tamilnadu, India
[c]Department of Physics, Sadakathullah Appa College (Affiliated to Manonmaniam Sundaranar University, Tirunelveli), Palayamkottai Tirunelveli, Tamilnadu, India

Article history
Received
27 June 2022
Received in revised form
06 January 2023
Accepted
29 January 2023
Published online
31 May 2023

*Corresponding author
spkpramu@sadakath.ac.in

Graphical abstract



Abstract

Tremendous success of machine learning algorithms at pattern recognition creates interest in new inventions. Machine learning in an era of big data is that significant hierarchical relationships within the data can be discovered algorithmically than other handcraft like features. In this study, Convolutional Neural Network (CNN) is used as feature descriptors in pulmonary malignancy prediction. Various feature descriptors such as Histogram of Oriented Gradient (HOG), Extended Histogram of Oriented Gradient (EXHOG) and Linear Binary Pattern (LBP) descriptors are analyzed with classifiers such as Random Forest (RF), Decision Tree (DT), K-Nearest Neighbour (KNN) and Support Vector Machine (SVM) for Computed Tomography (CT) The phenotype features of pulmonary nodules are important cues for identification. The nodule solidity is an important cue for white blob area identification. The method is analyzed in Lung Image Database Consortium (LIDC) dataset. Receivers Operating Characteristics (ROC) curves show the graphical summaries of detectors performance. It is proved that CNN based feature extraction with SVM classifier works well in pulmonary malignancy prediction.

Keywords: CT Images, Features, Descriptors, Classifiers, CNN, ROC

1.0 INTRODUCTION

Computed Tomography [1-5] is the best high resolution and volumetric, modality used to detect the characteristics of pulmonary nodule [6-8] in bronchial trees. Instead some of the nodules are in low quality resolution which the system could not identify itself or less experienced radiologists. Hence Computer Aided Design [9] improves the identification of nodule in previous studies. Annual Screening is recommended by doctors and associations for pulmonary cancer patients [10]. Early detection prevents the life of the patient, suggested by American society of clinical oncologies. Pulmonary nodules are grouped in to Elliptical, Lobulated, Spiculated and Spherical. The nodule named justaplura placed near the lung wall is dangerous to diagnose in segmentation process.

Various feature extractors and classification models are studied in [11]. The wavelet filters are useful to extract texture feature and have determined the number of corals using LBP descriptor because they use the pixel values from eight directions [12-15]. Deep feed forward Artificial Neural Network (ANN) [16-17] analyses the visual imagery. ANNs consist of a method of solving problems related to science through simple models that mimic the human brain, including behavior. An ANN is armed by small modules which simulate operation of a neuron. In ANN, only minimum number of layers is used. Deep feed-forward ANN analyses the visual imagery. The convolutional neural network has hundreds of hidden layers, which uses filters to extract features.

In Text Sentiment Classification, Recurrent Neural Network (RNN) network [12] is highly preferable where RNN uses the loop to translate sequence to sequence. Neurons within the same layer are not connected. So in RNN, the neurons in the layers are connected. RNN is the best for Reinforcement learning. Imagenet is suitable for Visual recognition challenges. Instead of autoencoder, we can use Conditional Random Field (CRF) model or deep sequential model. Handwritten digit recognition using backpropagation network. CNN was introduced by Lecun et al., in 1998. CNN allows multiple features to be extracted at each hidden layer. Convolutional Neural Network is used to classify tumors seen in lung cancer [18-19] screening. It has special properties such as spatial invariance and allows to multiple feature extraction. The deep CNN has been used widely since the accuracy of prediction increases dramatically. Lecun et al., have written a paper that CNN has been shown to eliminate the necessity of handcrafted feature extractors in gradient-based learning applied to document recognition [20]. Hence, multiple characteristics can be extracted and the learning process assigns weights appropriately to significant features [21].

The author Setio [22] showed the multi-view ConvNet. This is very much suited for false positive reduction and achieved good results for the nodule detection task in lung CT images. CNN consists of built-in preprocessing architecture [5]. A dominant dictionary like texture descriptor, texton is proposed as a feature [23]. Deep learning has proved a popular and powerful method in many medical imaging areas. In [22], CNN, Deep Neural Networks (DNN), and autoencoders are designed for lung cancer calcification.

Shape, Texture features [24] with Convolution patterns recognize the nodule area with less time and best accuracy value. Deep Learning ConvNet technique is well-matched for false positive reduction of a Computer Aided Design (CAD) system [25]. The low computation time of ConvNet is a decision aid in lung cancer screening [25]. The proposed framework, in order to understand the lung nodule features, heterogeneous computational feature type derived from the CNN as well as general low level HOG [26], EXHOG [27] and LBP features, exactly detect the nodules in CT pulmonary images. Furthermore extracted features from hybridized descriptors are classified using excellent classifiers SVM, KNN, DT and RF.

In recent year, we developed some methods for pulmonary node detection [28-29]. This research focuses on various combinations of features and classifiers on pulmonary malignancy detection. The results are analyzed using ROC curve.

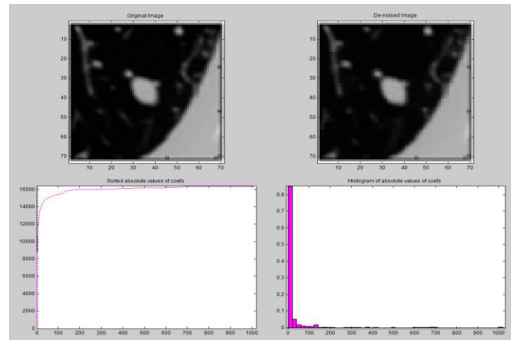
The remaining of the section is organized as follows: Section 2 elaborates the proposed method with all feature extraction methods and classifiers. Section 3 analyzes hybrid feature extractors and classifiers with conclusion in Section 4.

2.0 METHODOLOGY

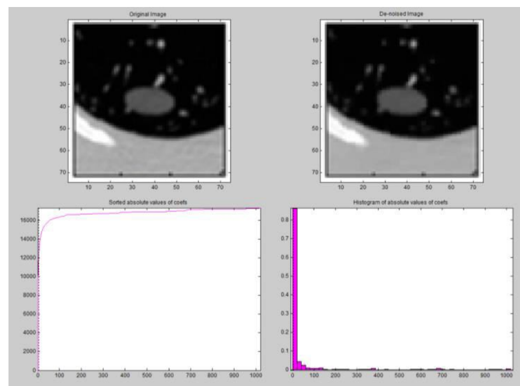
In this work, pre-processing, segmentation, feature extraction and classification are the important steps to detect nodule area.

2.1 Pre-processing

Pre-processing is essential prerequisite process. Noised images are not clear to surgery level. With filtering method, denoised image has been produced. Figure 1 shows the denoised image of ellipsoidal nodule.



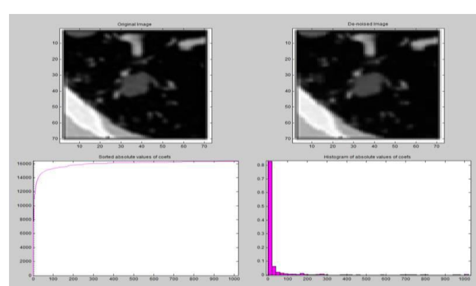
(a)



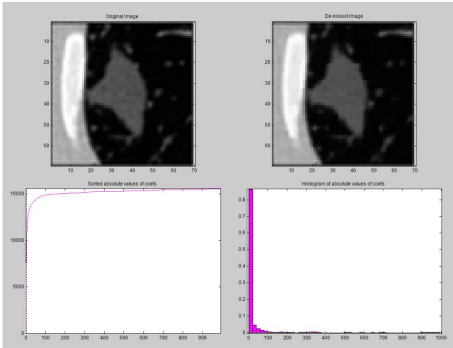
(b)

Figure 1 (a) Denoised image of Ellipsoidal Nodule (b) After rotation

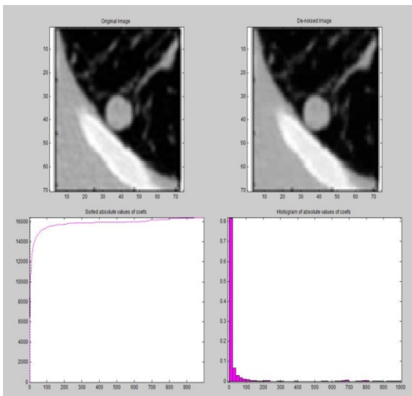
The above Fig. 1(a) shows the denoised image using wavelet Haar level 2 and Entropy as shanon. The thresholding method is soft threshold fixed form and global threshold value is 4.334. Number of bins value gives 50. In Figure 1 (b), there is an observation between rotated and nonrotated image in histogram. In histogram the values fall in to bins is raised above hundred where as the number of bins in original image is less than 100. Here the x – axis shows number of bins and y-axis shows the number of pixels. This is applied to four types of nodules which are shown in Figure 2.



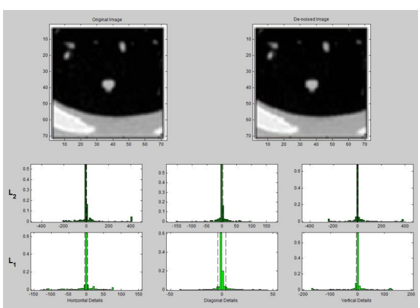
(a)



(b)



(c)



(d)

Figure 2 Original and Denoised image of (a) Lobulated Nodule (b) Spiculated Nodule (c) Spherical Nodule

Figure 2 (a-c) shows the sorted absolute values of coefts and histogram of absolute value of coefts. Here the absolute histogram gives clear idea than sorted histogram. Here the x – axis shows number of bins and y-axis shows the number of pixels. In Figure 2 (d), multi threshold value is applied to above spericity nodule. Using horizontal, vertical, diagonal details nodule location identified easily.

2.2 Segmentation

We use edge base segmentation in this work. The threshold value 128 is suitable to identify the edges during segmentation process. The results obtained after segmentation are shown in Figure 3.

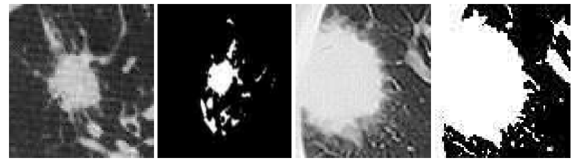


Figure 3 Edge based segmentation

2.3 Feature Extraction

2.3.1 Feature Extraction by Curvelet and Convolution Matrix

The merit of Curvelet is faster minimal redundancy. Fast Fourier Transform (FFT) is applied over input images. The Fourier frequency plane is sliced into circular and angular wedges and they are further decomposed into specific scales and angles respectively as shown in Figure 4.

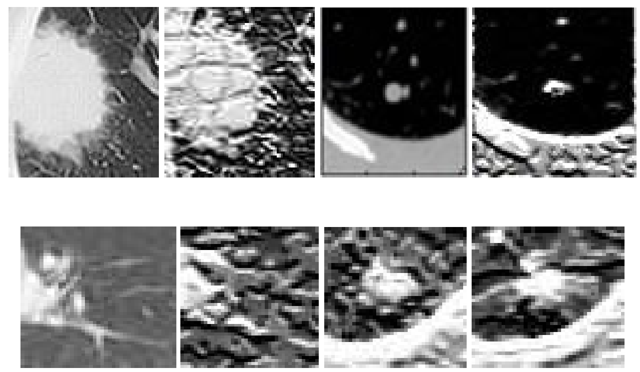


Figure 4 Nodules with Different Textures

The curvelet coefficients can be obtained by applying Inverse FFT (IFFT) over the wedges. Additionally, the curvelets can handle the edge discontinuities in a better way. The features of the segmented image are extracted by means of curvelet, which is known for its discriminative power. Fast Fourier Transform decomposes the nodule image in to frequency components which is shown in Figure 5.

This is equal to number of pixels in the image or spatial domain. The inverter transform rework the frequencies to image. In Figure 5, the first image is indeterminant and the last image shows determinant part. Convolution is one of the most important operations in nodule identification. This is mostly used in feature extraction and is also part of the block of CNN.

2.4 Classification

Classification is the separation of objects or data into different categories. When the classes are formed without the knowledge of data then the classes are called as priori classification. If the classes are created by features, then the classification is called as posterior classification. In classification the classes have been seemed a priori and consist of training the system so that when a new object is presented to the trained system, it will be able to assign the article to one of the prevailing categories.

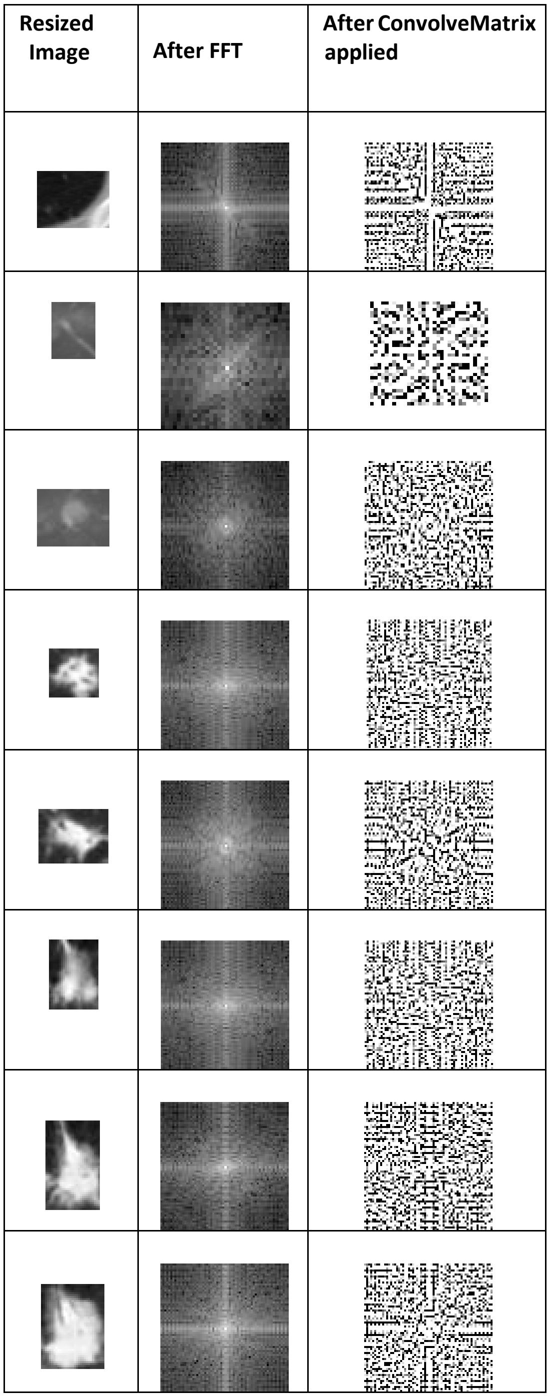


Figure 5 FFT and Convolution of input images

This point is called as supervised learning. For unsupervised classification [30] the classes are determined by the given data. But in supervised learning, one which provides a category label for objects is referred as training set. In unsupervised learning or clustering there is no explicit and the system forms clusters or natural groupings [31] of the input patterns [32]. Data mining techniques are used to find the nontechnical loss using classification algorithm.

SVM classifier

SVM [33, 34] is a popular and promising classifier that is trained with the malignant and benign CT images by means of the calculated threshold. Consider {1, 2, 3, …, N} CT images, which are to be classified as malignant and benign. Both these classes are divided by a hyperplane, which derives a criterion to separate the classes. Hence, the choice of hyperplane must be optimal, as the efficiency of the classifier depends on this hyperplane.

K-NN classifier

K-NN is the basic classifier which learns the nature of the training data, so as to differentiate between the normal and abnormal areas of the image. This classifier calculates the Euclidean distance between the image pixels and is denoted by

$$Dis_m = \sum_{i=1}^{N} \sqrt{x_i^2 - y_i^2} \quad (1)$$

The performance of the K-NN classifier is decided by the value of k and hence, k must be chosen carefully. However, manual choice of k is difficult and inefficient, as it requires prior knowledge about the dataset. Additionally, arriving at an optimal value of k by manual approach consumes more time and energy. In order to avoid this issue, an automatic k fold cross-validation scheme is employed, which can choose the value of k. The k fold cross-validation scheme works by decomposing the training images into k parts and a single part is considered as the testing and the rest of the images are treated as training objects.

This process is repeated until all the images are processed as test samples. At last, the mean value is computed for all the k results attained so far and this value is assigned as k. Hence, the value of k is assigned automatically, which is effortless and there is no need to have prior knowledge about the dataset.

Decision Tree Classifier

Generally, DT classifier does not suitable for simple classification. In order to create a model that predicts the value of target using decision value and pruning, can use DT classifier. In fact the DT is a nonparametric supervised learning method.

Random Forest Classifier

When more number of decision trees is used, it together forms RF classifier. The main purpose is to improve the classification rate of pulmonary nodule that depends on the values of random vector for all trees in the forest.

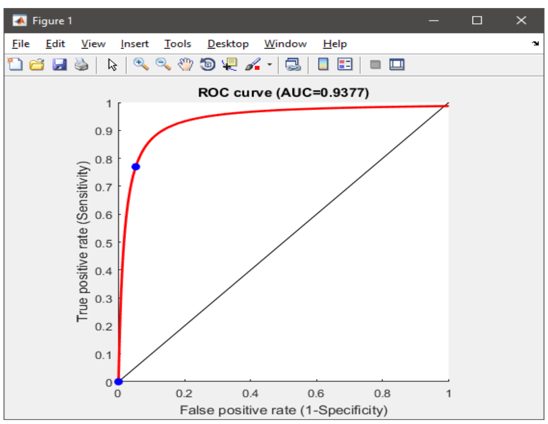
3.0 RESULTS AND DISCUSSION

The data set used for results analysis is LIDC IDRI [3] benchmark dataset which is available online. It consists of 1018 Pulmonary

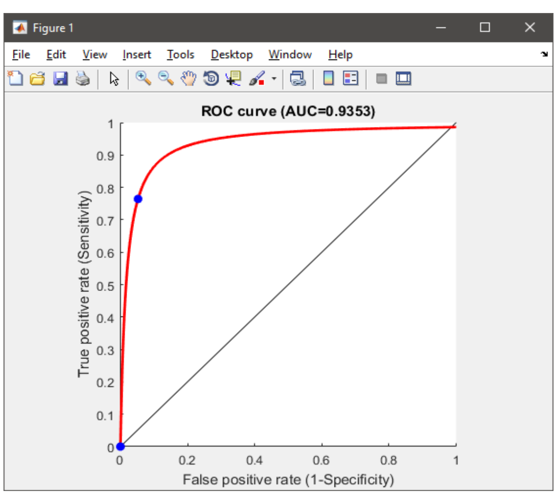
CT images that originated from a total of 1010 patients, in total 2,43,958 images. We analyze the proposed method with the help of ROC curve. ROC curve helps to evaluate the diagnostic performance of a test or the accuracy of a test to discriminate diseased cases from normal cases. The following combinations of feature descriptors and classifiers are used for analysis:

- SVM_HOG_CNN_LBP
- SVM_EXHOG_CNN_LBP
- DT_EXHOG_CNN_LBP
- DT_CNN_LBP
- DT_EXHOG
- DT_EXHOG_CNN
- KNN_CNN_LBP
- RF_CNN_LBP
- RF_EXHOG
- RF_EXHOG_CNN

Receivers Operating Characteristics (ROC) curves show graphical summaries of detectors performance. In early days, ROC curves are useful for analyzing detector performance of a radar or sonar detector. Here the ROC curve shows the performance of nodule detection. Figure 6 shows the ROC curve plot of SVM_HOG_CNN_LBP and SVM_EXHOG_CNN_LBP.



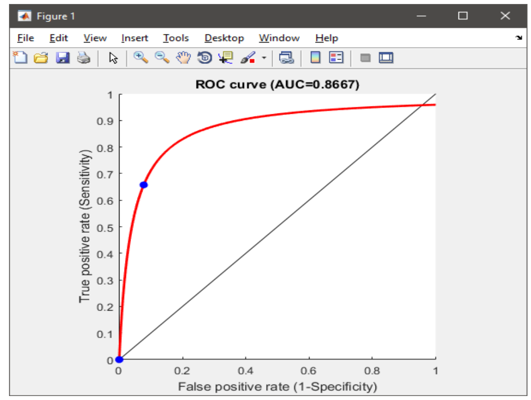
(a)



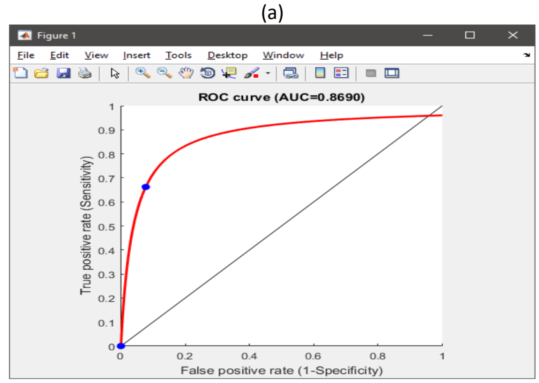
(b)

Figure 6 ROC Curve of (a) SVM_HOG_CNN_LBP (b) SVM_EXHOG_CNN_LBP

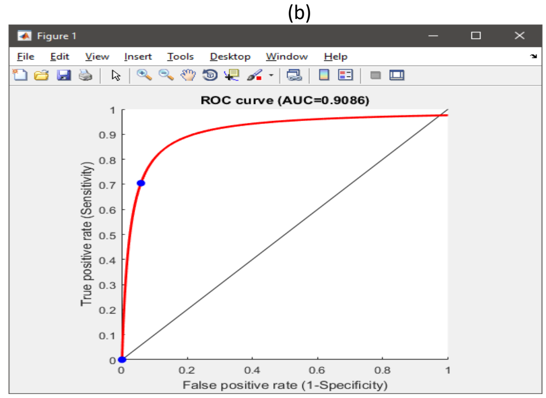
This curve illustrates the diagnostic capability of binary classification, with threshold value 128. Figure 7 displays ROC obtained using DT classifier.



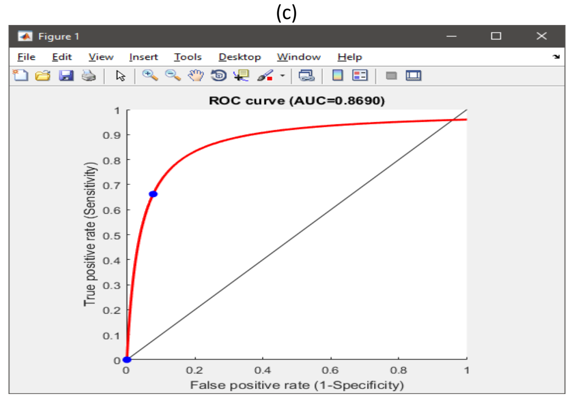
(a)



(b)



(c)



(d)

Figure 7 ROC Curve of (a) DT_EXHOG_CNN_LBP (b) DT_CNN_LBP (c) DT_EXHOG (d) DT_EXHOG_CNN

In Figure 7, red color line shows the True Positive Rate (TPR) suddenly increased up to the level of saturation for increasing False Positive Rate (FPR). Then after a certain value, it turns right side and it is slowly increased with the TPR. At this point, the x axis value in Figure 7 (d) is growing very little in percentage. This shows in the beginning stage, finding TPR is easy and after blue highlighted point or breaking point, finding TPR is a tougher one. Figures 8 and 9 shows ROC of KNN and RF classifier respectively.

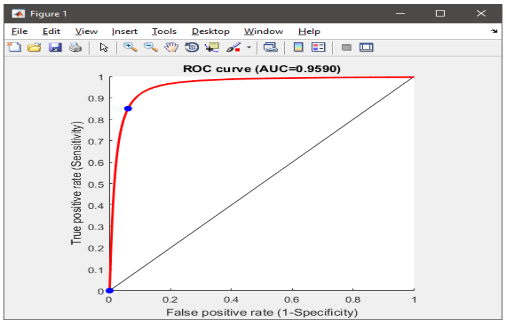
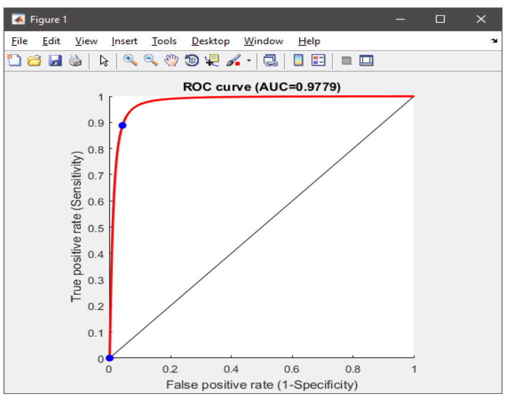
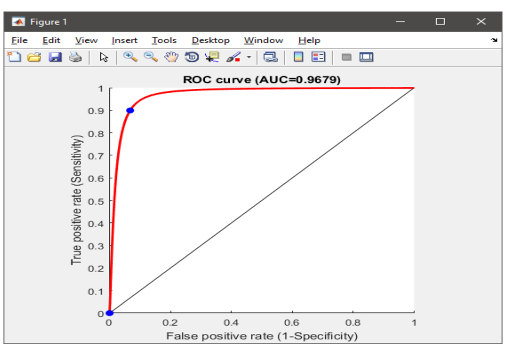


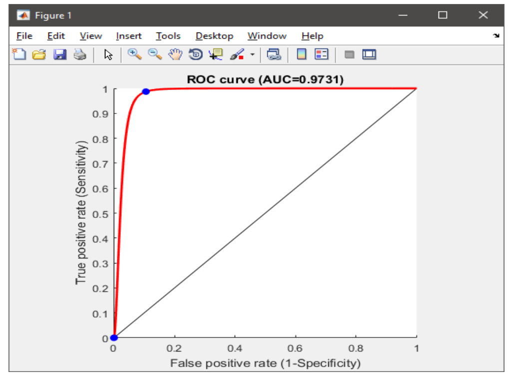
Figure 8 ROC Curve of KNN_CNN_LBP



(a)



(b)



(c)

Figure. 9 ROC Curve of (a) RF_CNN_LBP (b) RF_EXHOG (c) RF_EXHOG_CNN

ROC curves also plot the Probability of detection (Pd) versus the Probability of False Alarm (PFA). The probability of detection is one which says that the event occurred in RF_EXHOG in Figure 9 (b) and RF_EXHOG_CNN compositions in Figure 9(c).

If the result is zero, the event not occurred. In the case of nodule detection, '1' indicates that a target present and '0' indicates the nodule is not present. Area Under Curve (AUC) is used in classification analysis in order to determine which of the used models predicts the classes best. AUC is given as

$$\text{AUC} = \int_0^1 \text{TPR}(x)\,dx \tag{2}$$

Table I displays the values of AUC of all ROC curve analysis.

Table I Area Under Curve values for Decision Tree Classifier

Classifier	AUC
DT_CNN_LBP	0.86896
DT_EXHOG	0.90855
DT_EXHOG_CNN	0.86896
DT_EXHOG_CNN_LBP	0.86674
KNN_CNN_LBP	0.95905
RF_CNN_LBP	0.97786
RF_EXHOG	0.97313
RF_EXHOG_CNN	0.96786
SVM_HOG_CNN_LBP	0.93769
SVM_EXHOG_CNN_LBP	0.93527
SVM_EXHOG_CNN	0.9823

Here the TPRs are plotted against FPRs. The Closer AUC a model comes to 1, the better it is. So models with higher AUCs are preferred over those with lower AUCs. In Table I, AUC for the combination DT_EXHOG_CNN, DT_CNN_LBP are around 0.87. DT_EXHOG_CNN_LBP is the least value among other DT combinations. The result of DT is a 'good test' in Matlab 2019a for DT_EXHOG_CNN_LBP.

The AUC value of KNN classifier with CNN_LBP is 0.959 which shows better classification than DT. The comment received from Matlab 2019a is 'excellent test'. RF is the best classifier in terms of AUC with the combination of EXHOG, EXHOG_CNN & CNN_LBP. The value is above 0.96, which is also 'excellent test'.

The AUC of RF is quite excellent compared to DT and KNN Classifier. However, SVM is averagely best in all features classification as it achieves 0.98 value for nodule classification in Pulmonary CT images. Here By changing hyperplane distance or kernel value, the accuracy reaches the maximum. The AUC value of SVM_HOG_CNN_LBP is 0.93769. Standard Error value shows that 0.019 is negligible value. In this study, Area under the ROC curve lies with 95% confidence. In result of ROC Out of eleven, eight results are more than 90% in Confidence Intervals (C.I).

ROC curves can also be used to compare the diagnostic performance of two or more laboratory or diagnostic tests with Private Labs. ROC curves are given for four different classifiers with four different feature vectors. The curve is drawn between TPR and FPR and results are useful and easy to analyse the nodule types in bronchial trees of CT images.

4.0 CONCLUSION

This paper proposed a hybrid feature descriptor to find pulmonary nodule classification in deep learning environment. The feature vectors CNN, LBP, HOG and EXHOG are hybridized with machine learning classifiers SVM, DT, RF and KNN. The AUC evaluation on LIDC/IDRI dataset shows that SVM_EXHOG_CNN got better accuracy of 0.9823 which is comparatively higher than RF_EXHOG_CNN and DT_EXHOG_CNN. In future this can be extended with 3DCNN architecture using other classifiers.

Acknowledgement

The authors are thankful to Dr.S.M.Abdul Kader, Principal, Dr.M.Sheik Muhideen Badhusha, Dean of research and development, Sadakathullah Appa College, Tirunelveli for providing the necessary facilities in preparation of the paper.

References

[1] Abdulrazzaq, M. M., S. A. Noah, and M. A. Fadhil. 2015. X-Ray Medical Image Classification Based on Multi Classifiers. *2015 4th International Conference on Advanced Computer Science Applications and Technologies (ACSAT)*. 218-223. doi: 10.1109/ACSAT.2015.45.

[2] Setio, A.A.A., F. Ciompi, G. Litjens, P. Gerke, C. Jacobs, S. J. Van Riel, M. M. W. Wille, M. Naqibullah, C. I. Sánchez, and B. Van Ginneken. 2016. Pulmonary nodule detection in CT images: false positive reduction using multi-view convolutional networks. *IEEE transactions on medical imaging*. *35*(5): 1160-1169. doi: 10.1109/TMI.2016.2536809.

[3] Armato III, S.G., G. McLennan, L. Bidaut, M.F. McNitt-Gray, C. R. Meyer, A. P. Reeves, B. Zhao, D.R. Aberle, C. I. Henschke, E.A. Hoffman, and E.A. Kazerooni. 2011. The lung image database consortium (LIDC) and image database resource initiative (IDRI): a completed reference database of lung nodules on CT scans. *Medical physics*. *38*(2): 915-931. doi: 10.1118/1.3528204. PMID: 21452728; PMCID: PMC3041807.

[4] Niranjana, G., and M. Ponnavaikko. 2017. A Review on Image Processing Methods in Detecting Lung Cancer Using CT Images. *2017 International Conference on Technical Advancements in Computers and Communications (ICTACC).* 18-25. doi: 10.1109/ICTACC.2017.16.

[5] Chen, S., J. Qin, X. Ji, B. Lei, T. Wang, D. Ni, and J. Z. Cheng. 2016. Automatic scoring of multiple semantic attributes with multi-task feature leverage: a study on pulmonary nodules in CT images. *IEEE transactions on medical imaging*. *36*(3): 802-814. doi: 10.1109/TMI.2016.2629462.

[6] Gould, M.K., J. Donington, W. R. Lynch, P.J. Mazzone, D. E. Midthun, D. P. Naidich, and R. S. Wiener. 2013. Evaluation of individuals with pulmonary nodules: when is it lung cancer? Diagnosis and management of lung cancer, 3rd ed: *American College of Chest Physicians evidence-based clinical practice guidelines.* Chest. 143(5 Suppl):e93S-e120S. doi: 10.1378/chest.12-2351.

[7] Wang, H., X. H. Guo, Z. W. Jia, H. K. Li, Z. G. Liang, K. C. Li, and Q. He. 2010. Multilevel binomial logistic prediction model for malignant pulmonary nodules based on texture features of CT image. *Eur J Radiol*. 74(1): 124-9. doi: 10.1016/j.ejrad.2009.01.024.

[8] Naidich, D. P, A. A. Bankier, H. MacMahon, Schaefer-Prokop CM, Pistolesi M, J. M. Goo, P. Macchiarini, J. D. Crapo, C. J. Herold, J. H. Austin, and W. D. Travis. 2013. Recommendations for the management of subsolid pulmonary nodules detected at CT: a statement from the Fleischner Society. *Radiology.* 266(1): 304-17. doi: 10.1148/radiol.12120628.

[9] Gonçalves, L., J. Novo, and A. Campilho. 2016. Hessian based approaches for 3D lung nodule segmentation. *Expert Systems with Applications*. 61: 1-15. https://doi.org/10.1016/j.eswa.2016.05.024

[10] Bhuvaneswari, P., and A. Brintha Therese. 2015. Detection of Cancer in Lung with K-NN Classification Using Genetic Algorithm. *Procedia Materials Science*. 10: 433-440. https://doi.org/10.1016/j.mspro.2015.06.077.

[11] Sowmyayani, S., 2022. Rapid Study on Feature Extraction and Classification Models in Healthcare Applications. *International Journal of Computer and Systems Engineering.* 16(12): 620-627.

[12] Nurtiyasari, D., D. Rosadi, and Abdurakhman. 2017. The application of Wavelet Recurrent Neural Network for lung cancer classification. *2017 3rd International Conference on Science and Technology - Computer (ICST),* 127-130. doi: 10.1109/ICSTC.2017.8011865.

[13] Epifanio, I., and G. Ayala. 2002. A random set view of texture classification. *IEEE Transactions on Image Processing*. 11(8): 859-867. doi: 10.1109/TIP.2002.801119.

[14] Soille, P. 2004. Texture Analysis. In: *Morphological Image Analysis*. Springer. Berlin, Heidelberg. DOI: 10.1007/978-3-662-05088-0-11.

[15] Ciompi, F, C. Jacobs, E.T. Scholten, M.M. Wille, P. A. de Jong, M. Prokop, and B. van Ginneken. 2015. Bag-of-frequencies: a descriptor of pulmonary nodules in computed tomography images. *IEEE Trans Med Imaging*. 34(4): 962-73. doi: 10.1109/TMI.2014.2371821.

[16] Amirani, M.C., Zahra Sadeghi Gol, and Ali Asghar Beheshti Shirazi. 2008. Efficient Feature Extraction for Shape-Based Image Retrieval. *Journal of Applied Sciences. 8: 2378-2386.* **DOI:** 10.3923/jas.2008.

[17] Rao, P., N. A. Pereira, and R. Srinivasan. 2016. Convolutional neural networks for lung cancer screening in computed tomography (CT) scans. *2016 2nd International Conference on Contemporary Computing and Informatics (IC3I)*. 489-493. doi: 10.1109/IC3I.2016.7918014.

[18] Song, Q., Lei Zhao, XingKe Luo, and XueChen Dou. 2017. Using Deep Learning for Classification of Lung Nodules on Computed Tomography Images. *Journal of Healthcare Engineering*. DOI: 10.1155/2017/831470.

[19] Dhaware, B. U., and A. C. Pise. 2016. Lung cancer detection using Bayasein classifier and FCM segmentation. *2016 International Conference on Automatic Control and Dynamic Optimization Techniques (ICACDOT).* 170-174. doi: 10.1109/ICACDOT.2016.7877572.

[20] Sammouda, R. 2016. Segmentation and Analysis of CT Chest Images for Early Lung Cancer Detection. *2016 Global Summit on Computer & Information Technology (GSCIT).* 120-126, doi: 10.1109/GSCIT.2016.29.

[21] Lecun, Y., L. Bottou, Y. Bengio, and P. Haffner. 1998. Gradient-based learning applied to document recognition. *Proceedings of the IEEE.* 86(11): 2278-2324. doi: 10.1109/5.726791.

[22] Dhara, A. K, S. Mukhopadhyay, A. Dutta, M. Garg, and N. Khandelwal. 2016. A Combination of Shape and Texture Features for Classification of Pulmonary Nodules in Lung CT Images. *J Digit Imaging.* 29(4): 466-75. doi: 10.1007/s10278-015-9857-6.

[23] Quratul, A., Jaffar, Arfan Choi, and Tae-Sun. 2014. Fuzzy anisotropic diffusion based segmentation and texture based ensemble classification of brain tumor. *Applied Soft Computing*. 21: 330–340. Doi: 10.1016/j.asoc.2014.03.019.

[24] Beijbom, O., P. J. Edmunds, D. I. Kline, B. G. Mitchell, and D. Kriegman. 2012. Automated annotation of coral reef survey images. *IEEE Conference on Computer Vision and Pattern Recognition.* 1170-1177. doi: 10.1109/CVPR.2012.6247798.

[25] Pun, C. M., Zhu, and Hong-Min. 2009. Textural image segmentation using discrete cosine transform. *Proceedings of the 3rd International Conference on Communications and Information Technology*. pp. 54-58.

[26] Goharian, and Grossman. 2003. Data Mining Classification, Illinois Institute of Technology. Chidanand Apté, Sholom Weiss, Data mining with decision trees and decision rules. *Future Generation Computer Systems*. 13(2–3): 197-210, https://doi.org/10.1016/S0167-739X(97)00021-6.

[27] Satpathy, A., X. Jiang, and H-L. Eng. 2011. Extended Histogram of Gradients with Asymmetric Principal Component and Discriminant Analyses for Human Detection. *Proceedings - 2011 Canadian Conference on Computer and Robot Vision, CRV*. 64-71. 10.1109/CRV.2011.16.

[28] Sathik, M.M., and S. P. Kailasam. Pulmonary CT Image analysis for Nodule Detection using Inspired FCM Clustering.

[29] Kailasam, S.P., and M.M. Sathik. 2019. A novel hybrid feature extraction model for classification on pulmonary nodules. *Asian Pacific Journal of Cancer Prevention.* 20(2): 457.

[30] Lu, D., and Q. Weng. 2007. A survey of image classification methods and techniques for improving classification performance. *International Journal of Remote Sensing.* 28(5): 823 870. DOI: 10.1080/01431160600746456

[31] Naseem, R., K. S. Alimgeer, and T. Bashir. 2017. Recent trends in Computer Aided diagnosis of lung nodules in thorax CT scans. *2017 International Conference on Innovations in Electrical Engineering and Computational Technologies (ICIEECT)*. 1-12, doi: 10.1109/ICIEECT.2017.7916548.

[32] Sherly Alphonse A., and Dejey Dharma. 2018. Novel directional patterns and a Generalized Supervised Dimension Reduction System (GSDRS) for facial emotion recognition. *Multimedia Tools Appl*. 77(8): 9455–9488. https://doi.org/10.1007/s11042-017-5141-8

[33] Gopi, K., and J. Selvakumar. 2017. Lung tumor area recognition and classification using EK-mean clustering and SVM. *2017 International Conference on Nextgen Electronic Technologies: Silicon to Software (ICNETS2).* pp. 97-100. doi: 10.1109/ICNETS2.2017.8067906.

[34] Sankari, E.S., and D. Manimegalai. 2017. Predicting membrane protein types using various decision tree classifiers based on various modes of general PseAAC for imbalanced datasets. *J Theor Biol.* 21(435): 208-217. doi: 10.1016/j.jtbi.2017.09.018