Lung Nodule & IPMN Characterization using GIST Features based on Modified SVM

DR G.BASKAR[1] , R.SUBALAKSHMI[2]

Abstract :

Utilizing portable computer-assisted diagnosing (CAD) tools, tumor characterization from radiology pictures is accurate and quicker. Neoplasm characterization exploitation tools also provide cancer staging, analysis, and personal treatment as exactness medication. We plan new, improved machine learning algorithms during this study to induce higher tumor characterization. We tend to plan two differing types of tumor characterization followed in my research. The primary approach employs a 3D Convolutional Neural Network and Transfer Learning. Driven by radiologists' interpretations of scans and structured project features stand for into a CAD system via Multi-Task Learning (MTL) framework with modified SVM with instance weighting algorithm. Therefore the second style of approach has used GIST based on tumor characterization by using the cluster with modified SVM algorithm relating to that to handle the restricted accessibility of classified education knowledge, a standard-issue in scientific imaging applications. We were impressed by learning from label share (LLP) processes in pc vision. We applied modified SVM with instance weighting to categorize tumors and evaluated our proposed supervised learning algorithms on two unique tumor diagnosis trials: lung nodule and pancreas (IPMN) with 1018 CT and 171 MRI severally. The experimental results incontestable that the accuracy of and lung nodule category was superior to existing algorithms.

Keywords: Machine Learning Algorithms (Modified SVM Algorithm), Lung nodule cancer, IPMN, Pancreatic cancer, GIST features.

I. INTRODUCTION

Cancer is showing an emotion-charged issue in our time; scores of individuals worldwide are afflicted with it, and there's still no cure. Taking management of it through early detection, on the opposite facet, could also be some way to increase the survival rate. Carcinoma is the second most typical cancer sort in each man and girl, following prostate and carcinoma. The Yankee

Cancer Society classified carcinoma as the foremost aggressive cancer in 2016 [1], with a death rate of quite seventy. As a result, if the cancer is detected early, once it's restricted to the respiratory

organ and has not unfolded to the liquid body substance nodes, the chance of survival will increase to forty-nine nothing [2,3].

The main aim of this research is to do advanced diagnosing and to expand care services. Eexploitation pc power-assisted diagnosing (CAD) within the medical workflow assists physicians in creating precise selections and improves diagnosing accuracy. Machine learning ways generate a coaching model for medical pictures and handle all information objects within the pc help structure. Numerous sophisticated image process techniques and segmentation methods are applied to pictures to classify tumors in ancient CAD diagnosing systems. As a result, extracting low-level options from pictures may be troublesome [4]. Pancreatic cancer (PC) may be an extremely malignant channel growth that poses important challenges in early detection and later treatment [5,6]. Consistent with the Yankee Cancer Society, the death rate of patients with laptops continues to rise, ranking fourth within the U.S. [7,8]. it's calculable that some fifty-seven,600 individuals are diagnosed with a laptop, and some forty-seven,050 individuals can die from laptops in 2020, creating laptop AN incurable illness [9,10]. PCs are still widely utilized in developing countries. Because of this, a comprehensive surgical laptop diagnostic is critical, particularly in the detection of laptop staging, which can help doctors deliver the most effective therapeutic schedule for various stages of laptop and patients to receive early medical interventions before advanced laptop kind [11,12]. Most cancers within the respiratory organ tissue, generally in air passages and line cells, are malignant neoplastic diseases. Respiratory organ illness is defined by AN uncontrolled cell growth in respiratory organ tissue that is power-assisted by a malignant respiratory organ growth. This growth may also unfold to encompassing tissue or alternative body elements outside the respiratory organ via metastasis [14]. some four-hundredth of individuals are diagnosed with cancer at some purpose in their lives, with a typical death rate of 171.2 per 100,000 individuals per annum. carcinoma is one of the foremost common diseases within the fitness trade, inflicting varied deaths with a series of diseases moving the cardiovascular system, together with non-small cell carcinoma, tiny cell carcinoma, respiratory organ tumor tumors, and respiratory organ failure [13]. Carcinoma is the focus of this study, which focuses on the troublesome drawback of machine-controlled prognosis of Intraductal outgrowth glycoprotein Neoplasm's (IPMN). As a pre-malignant disorder, IPMN may lead to the development of invasive cancer if not treated. Within the primary epithelial duct and its branches, IPMN may be a mucin-producing tumor Cancers of the lung and carcinoma are two of the most frequent [15]. Whereas carcinoma

is the foremost reason behind cancer-related deaths worldwide, carcinoma has the worst analysis, with a 5-year survival rate of solely seven-member within the U.S. [16]. Detection and characterization of that respiratory organ and exocrine gland tumors will facilitate early prognosis and, as a result, enhanced survival risk with acceptable treatment/surgical operation plans. They're radio graphically detectable precursors to the bulk of exocrine gland cancers [17]. The CAD structures are designed to reduce the amount of false positives and false negatives by allowing radiologists to make quick and accurate judgments. There are several advantages to having a diagnostic procedure that is more sensitive to false flags than there are to having a tumor mistakenly labeled as benign. In the literature review, computerized methods for detecting and analyzing cancers in specific organs, such as the breast, colon, mind, lung, liver, and prostate, are discussed. In addition to [18, 19, and 20] Deep learning approaches have been identified in the scientific image analysis literature, a change from engineering to feature learning. Convolution neural networks are used to extract features, and then Random Forest (RF) is used to classify the data [22, 23]. Radiologists rely on formalized summaries and automated imaging alternatives investigations to make their diagnostic decisions. The combination of two independent system learning algorithms is described in this work as one of the easiest ways for growth characterization from radiology images. IPMI 2017 [24] may benefit from this research, which focuses on a unique, monitored rule. It was observed that nodules detected in Computerized Tomography scans had a high probability of developing GIST, hence a new sophisticated supervised learning set of rules was devised to take this into account. Finally, a graph normalized framework has been used to provide the final malignancy result. In order to acquire a malignancy grade, respiratory nodule features are complete, and those qualities complement each other. Analysis of every home-made and deep learning capacity to determine how accurate they are when used to growth characterization of respiratory organ nodules and IPMNs is the primary objective of this study. In order to generate a first set of labels in an uncontrolled way, we prefer to group the samples into distinct agencies at intervals in the feature space. As part of this research, we used label proportions to train the Proportion-Support Vector Machine, which has been altered by the technique of changing SVM using the Instance Coefficient Rule. This study is expected to be the most accurate when it comes to determining the difference between malignant and benign tumors. Tumors were represented using two distinct types of machine learning methods that were both supported. It has been proposed to use supervised learning in the first kind of 3D CNN architecture that enabled

Graph regularization and performance assessments for IMPN classification from CT images. As a second sort, GIST options supported growth characterization by exploitation cluster with changed SVM rule, which gave better results than the current algorithms, for the classification of respiratory organ nodules from COMPUTED TOMOGRAPHY scans and exocrine gland cysts from MAGNETIC RESONANCE IMAGING scans. The categorization of the tumor is the subject of this research. Next, send us a thank-you message for our part in developing the strategy. In Section 3, we'll offer a novel supervised learning approach for identifying respiratory organ nodules and IPMN from CT and MRI images, respectively. This section contains all about the experiments and their results. Section five comes to an end with a statement of conclusions.

II. BACKGROUND OF THE STUDY

This section introduces the ideas utilized in this paper and suggests a changed machine learning formula. In this research, the outcome has planned a way of computer-aided identity.

This section introduces the ideas utilized in this paper and advises a changed machine learning formula. The goal of this research was to develop a method for detecting pulmonary nodules and IPMN cysts.

Deep learning has returned to virtually all areas of visible analysis in the Associate in nursing age in order to explore the impact of functions derived from particular deep learning architectures. This study, to the best of its ability, examines the automated analysis of IPMNs from MRI.

Rather than hard-assigning labels, the planned machine learning formula estimates label proportions supported information. Moreover, two completely different strategies are planned to mitigate the impact of droning labels (i.e., Mislabels) no inheritable throughout clustering: use developed SVM with Instance Weighting, trained on label proportions most effectively, and cluster with changed SVM, that is, characterization on GIST features.

2.1. *Supervised Learning Algorithm*

Consider X = [x1, x2 . . . xn] T X R $^{n \times d}$ correspond the information capabilities obtain from pictures of lung nodules, having dimension d. In this model, it has an attribute. The score has been calculated from y1,y2…N, assigned to Y. The functions extracted from images represent X, the 1-5 scale represents malignancy rating where one is benign, and five is malignant. Then the research

$$\min_{W} \|XW - Y\|_2^2, \quad \text{s.t.} \quad \|W\|_1 \leq t. \tag{1}$$

focus on training the labeled data for the supervised learning and focuses on the co-efficient vector along with the regression estimator where W is related to d. Y is focused on characterization. To avoid over fitting, a regularize is added. Then followed with the traditional least square to overcome the optimization problem. The formula to calculate the least square

Then followed with the sparsity coefficient vector w1,w2…wd, which is assigned to W, is carried out by the parameter t in the equation stated above. Functions exposed to constraints are minimized by convex sets and their corresponding constraints in equation I, which is convex. It has been shown that both supervised learning's using deep learning and MTL categories benign and malignant.

2.2 3D Convolutions Neural Network (CNN) and Fine-Tuning

The goal of this study is to analyze the visual features of a joint model. For lung nodule characterization, the Sports-1M data set was used. Classes and 1 million videos have been assembled for the Sports-1M data collection. Many samples with outstanding calibration have been obtained in this study to offset the absence of a lung nodule data collection. Sports-1M will provide you with a comprehensive feature representation. Convolution sets, absolutely-related layers, and a category layer compose the 3D CNN. The 3-D Convolution Neural Network entry's dimensions, where sixteen specifies the number of slices, are detected using a max-pooling layer. Because the image's constant dimensions have been changed, the data set now contains three channels and sixteen slices instead of the previously used eight. As a result, the total number of entrants can be computed, and sixteen slices have been generated. The first three convolution layers contain 64, 128 and 220 filters, while the latter two levels have 512 filters.

The length of vectors utilized as an entrance point into the Multi-task framework structure is four thousand and ninety-six, the dimensions of the fully-related layers.

Multiple tasks may be learned at the same time using the Multi-tasking Learning (MTL) approach. In light of M's obligations, I'm doing this study to find a model for a project that makes use of records. The study starts with M-task multi-task learning and ignores the shared function. The job includes parameters We and allocated to I, a square matrix, rank is used to indicate c cardinality, and nuclear trace norm is constructed using low-rank solutions [26].

2.2. *Optimization*

The traditional process uses the up-to-the-minute gradient descent as an optimization set of rules. Conversely, because the '1norm is not differentiable at W = zero, and the gradient descent move towards fails to offer sparse solutions [27], widespread gradient descent cannot be performed here. For this reason, we employ the increased proximal gradient approach [28]; the gradient method is to solve the convergence charge and control the number of iterations.

2.3. *Unsupervised learning methods*

The capability of unsupervised learning techniques for tumor characterization struggle will be investigated in the second part of this paper. There are three steps in our research, unsupervised framework. To start, we use clustering to estimate the set of labels based on the images. In this research, the labels are calculated to obtain label proportions from each cluster. To study the categorization of tumors using clusters and label proportions, the first estimation is done by assigning x1,x2…xn to X. R n×d consider in making the network such that matrix which is input incorporated images in such a way that the relation x belongs to d. To overcome noisy labels, we express the latent variables.

III. METHODOLOGY

In this segment, we illustrate our planned models in detail. Our proposed supervised learning algorithms (Modified SVM based on Instance Weighting & cluster with Modified SVM) were functional to identify tumor diagnosis trials: lung nodule and pancreas (IPMN) with one thousand and eighteen COMPUTERIZED TOMOGRAPHY and one seventy-one MAGNETIC RESONANCE IMAGING scans, along with GIST features. MATLAB software was used to estimate our proposed model.

3.1. *Proposed Model 1: Modified SVM Algorithm*

This research is to start with the reduction techniques [29]. The sampling states that the estimator to consider the number of instances. This work presents the weighting set of rules for variance reduction. The algorithm used in this work is feature selections on the vector principles. [30], [31] to examine the functions of vectors and the concept used is instance weighting. The instance within the real function area can be represented by a new vector inside the vector function area described below by decomposing an example vector along each dimension.

Definition 1: Assume x = (x1,..., xd) is an example within the genuine feature area d, and x^H and x^M are the nearest times to x with the same and opposite elegance labels, correspondingly. For each x d, x may be mapped to x in a new characteristic space d according to:

$$x j = |xj − x^M j | − |xj − x^H j | 2$$

Xj is the jth coordinate of x in the new feature space d, and xj , x^M j , or x^H j is the jth coordinate of x, $x^{H,}$ or x^M in d, respectively. Vector x is the vector of x, and d is the vector feature space—instance Weighting Algorithm.

The preceding definition and example of vector characteristic space best take one nearest neighbor from each class into account. To decrease the effect of noise or outliers in the training set on the converted function area, an example vector may be computed using a couple of nearest neighbors from each magnificence. All of our neighbors from each class for a specific example of Eq (3), as follows:

Method 1:Proposed Model for Modified SVM based on Instance Weighting Algorithm where x^{Hl} j or x^{Ml} j means the jth element of the Ith neighbor to x with the same or specific class labels,

respectively. M or h characterizes the full quantity of misses or hits (m + h equals the total number of instances where x^{Hl} j or x^{Ml} j represents the jth in the training set excluding the given instance). Once the vector function area has been generated, the next task is to maximize the disparity of vectors in this area to weight instances within the unique space. To calculate the outlying degree of any vector x, we calculate its average distance from all other vectors; a greater average distance indicates a higher outlying degree.

3.2. *Proposed Model 2: Cluster with Modified SVM Algorithm*

This paper will evaluate the opportunity of unsupervised learning approaches for tumor characterization problems. As shown in Figure 1, our anticipated unsupervised framework consists of three steps. To begin, we use clustering to estimate an initial set of labels based on the appearance features extracted from the images. Afterward, we work out label proportions for each cluster using the initial labels obtained. At last, we study tumor categorization using the initial cluster assignments and label proportions. Table I contains a list of various supervised and unsupervised learning experiments and their evaluation sets.
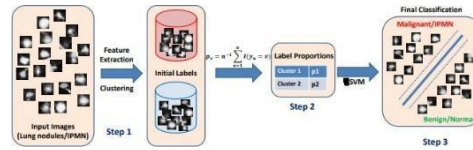
Figure 1: An outline of the second method of the proposed unsupervised approach.

TABLE I: machine learning along with their evaluation sets.

No. of experiments	Details	Evaluation Set
E1	3D CNN based multi-task learning with attributes, fine-tuning (C3D) network, Modified SVM with instance weighting algorithm	3D dataset: Malignancy score regression of Lung nodules (CT) 3D dataset: Lung nodules (CT) and IPMN classification (MRI)
E2	GIST features, Cluster with Modified-SVM	3D dataset: IPMN classification (MRI)

IV. EXPERIMENTS RESULTS

4.1 Data for Lung Nodules

The research used the LIDC-IDRI dataset from the Lung Image Database Consortium [32] for testing and evaluation, one of the widely existing lung nodule datasets. The dataset comprises 1018 CT scans with slice thicknesses ranging from 0.45 mm to 5.0 mm. Four radiologists, as a rule, annotated lung nodules with diameters equal to or greater than 3.0 mm. We looked at nodules that at least three radiologists had interpreted for evaluations. 1340 nodules met this condition. Since a nodule's malignancy and attribute scores may differ from radiologists, their mean scores were used. There were 509 malignant and 635 benign nodules in the concluding evaluation set. The images were resample to be isotropic, with 0.5 mm spacing in each dimension, as a pre-processing step.

4.2 Data for IPMN

T2 MRI axial scans from 171 subjects are included in the data for IPMN classification. A radiologist classified the scans as normal or IPMN. There were 38 normal subjects among the 171 scans, while 133 were from IPMN patients. The scan's in-plane spacing (xy-plane) ranged from 0.468 mm to1.406 mm. As a pre-processing component, we concern N4 bias field correction [35] to each picture to optimize variations in image intensity. The image is then smoothed while the edges are conserved using a curvature anisotropic image filter. Table I contains a list of various supervised and unsupervised learning experiments and their evaluation sets.

4.3 Experiment 1

The Sports1M dataset [34] with its 487 classes was used to fine-tune the 3D CNN network. We used the midpoint as a pivot point to train the network with binary labels for malignancy and the six traits, and tagged samples as positive (or negative) based on their scores being larger than the midway (or less than). In this context, malignancy and its characteristics are referred to as "tasks." The C3D was fine-tuned and 10-fold cross-validated using these seven activities. In addition to fine-tuning the network, we reduced the need for a large quantity of labeled training data. Because the network requires three-channel picture sequences with at least 16 slices, we concatenated the grey level axial channel with the other two. The interpolation method was used in order to ensure that all of the input volumes had 16 slices. The 4096-dimension 3D CNN's first completely linked layer produced the final feature representation.

Our suggested method was evaluated using both classification and regression measures. If the projected score of a nodule is within 1 of the actual score, we consider the nodule to have been correctly categorized. For regression, we determined the average absolute difference in predicted and actual scores. Our suggested MTL technique is compared in Table II to approaches like GIST features [36], 3D CNN features from a trained network + LASSO, Ridge Regression (RR), and 3D CNN MTL+trace norm. Table II In terms of classification accuracy and mean score difference, our suggested graph-regularized MTL surpasses existing techniques. There was a 25% increase in classification accuracy for GIST and a 21% increase in classification accuracy for trace norm. The suggested MTL enhances performance by 5% over the pre-trained network. For GIST and trace-norm, our technique lowers the average absolute score difference by 42% and 37% respectively.

4.4 Experiment 2

Both lung nodules and IPMN datasets were used for unsupervised learning evaluations. To calculate image-level features, we make use of GIST descriptors [36]. For both benign and malignant tumors, the number of clusters is limited to two. An analysis was conducted to establish whether the clusters were benign or cancerous. We conducted a 10-fold cross-validation to test

our suggested method. A linear-kernel SVM was trained using the clustering-generated label proportions and training data.

To evaluate the effectiveness of our unsupervised technique, we looked at its accuracy, sensitivity, and specificity. On the basis of accuracy and sensitivity, the proposed permutation of clustering and modified SVM is superior to alternative methods (see Table III). In contrast to clustering+Modified SVM, the second suggested framework enhances sensitivity for lung nodules by about 31 percent and IPMN classification by almost 17 percent . In contrast to the suggested strategy, which assigns a disproportionate number of examples as benign (normal), the clustering, clustering+modified SVM, and clustering+RF approaches have high specificity but poor sensitivity. When comparing lung nodules and IPMN, the proposed method's accuracy is up 34% and 19% higher than clustering.

TABLE II: meanabsolute score difference for lung nodule characterization.

| Methods	Accuracy (%)	MeanScore Difference
GIST features + LASSO	76.83	0.675
GIST features+ RR	76.48	0.674
3D CNN features + LASSO (Pre-trained)	86.02	0.530
3D CNN features+ RR (pre-trained)	82.88	0.597
3D features | LASSO	88.04	0.497
3D CNN features+ RR(Fine-tuned)	84.53	0.550
3D CNN MTL with Trace norm	80.08	0.626
3D CNN with Multi-task	91.26	0.459
Our Proposed 3D CNN with MLTFF	96.29	0.326

TABLE III: Classification of IPMN using GIST feature and supervised learning classifiers.
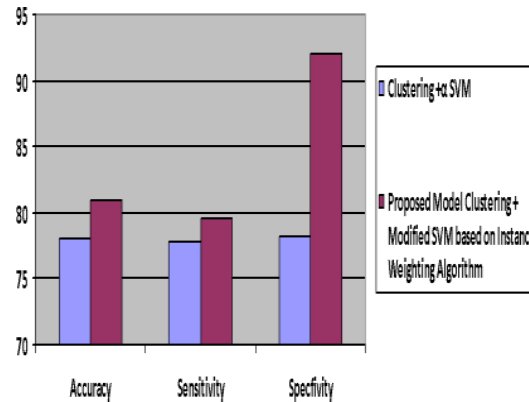
TABLE IV: IPMN classification with GIST features.

Features	Evaluation Set	Classifiers	Accuracy (%)	Sensitivity (%)	Specificity (%)
	Lung Nodule Classification	Clustering +α SVM	78.06%	77.85%	78.28%
		Proposed Model 1 ⟶ Clustering + Modified SVM with Instance Weighting Algorithm	80.91%	79.63%	92.09 %
	IPMN classification	Clustering+αSVM	58.04%	58.61%	41.67 %
		Proposed Model 1 ⟶ Clustering +Modified SVM with Instance Weighting Algorithm	64.81%	62.90%	66.1 %

| Features	Classifiers	Accuracy (%)	Sensitivity (%)	Specificity (%)
GIST	Clustering +SVM	76.05%	83.65%	52.67%
	Clustering +RF	81.9%	93.69%	43.0%
	Proposed Model 2	83.07%	93.73%	54.32%

The below Figure 2: results for IPMN.
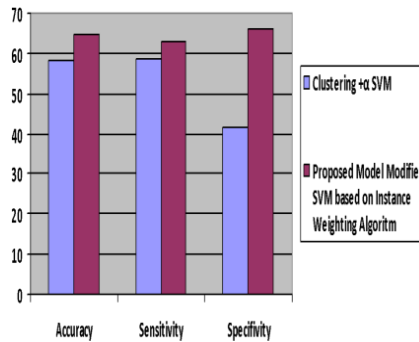
A) IPMN Classification Clustering + Modified SVM

b) Lung Nodule Classification Clustering + Modified SVM

The following Figure 3: 1(c),2 (d) (With GIST feature) in Modified SVM based on Instance Weighting Algorithm & Clustering with Modified SVM.
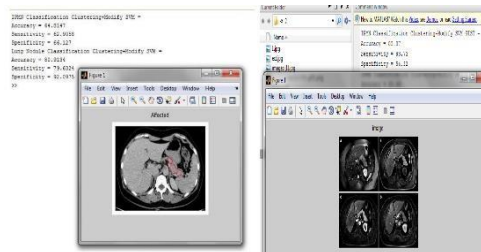
(c) IPMN classification Clustering+Modified SVM with Instance weighting algorithm   (d) IPMN classification clustering+Modified SVM

V. CONCLUSION

Using 3D CNN-based graph regularized sparse MTL and GIST features to determine the malignancy of lung nodules, we provide two new frameworks in this research for doing so. MTL and transfer learning for 3D deep networks may be used to improve the risk classification of lung nodules in this research. Various pre-trained deep networks are under investigation for their effect on lung nodule categorization. We've found that commonly used imaging alternatives like GIST provide results that are on par with those produced by pre-trained network features. ImageNet classification tasks are often used to train deep neural networks. The CNN filters were adjusted to the subtleties of natural images rather than medical images. To the best of our knowledge, this is one of the earliest and biggest assessments of an IPMN classification CAD system. As a relatively new area of study, IPMN categorization using CAD systems necessitates the investigation of several imaging modalities. Since CT images have a greater resolution than MRI images, they are often utilized in conjunction with MRI as an additional imaging modality for diagnosing pancreatic

cysts. Detection and segmentation of the pancreas may also be used to provide more accurate predictions about IPMN and cysts. The architecture of the pancreas, especially in MRI images, may make it difficult to phase the organ.. The use of various imaging modalities for joint segmentation and the detection of pancreatic cysts and IPMN may be employed to address this problem. Unsupervised learning methods will definitely be researched in addition to those areas of medical imaging that have unique issues.

VI. REFERENCES

[1]. Firmino, m.; angelo, g.; morais, h.; dantas, mr; valentim, r. Computer aided detection (cade) and diagnosis (cadx) system for lung cancer with likelihood of malignancy. Biomed. Eng. Online 2016.

[2]. Wang, M.; Long, F.; Tang, F.; Jing, Y.; Wang, X.; Yao, L.; Ma, J.; Fei, Y.; Chen, L.; Wang, G.; et al. Autofluorescence Imaging and Spectroscopy of Human Lung Cancer. Appl. Sci. 2017.

[3]. Wingo, P.A.; Ries, L.A.G.; Giovino, G.A.; Miller, D.S.; Rosenberg, H.M.; Shopland, D.R.; Thun, M.J.; Edwards, B.K. Annual Report to the Nation on the Status of Cancer, 1973–1996, With a Special Section on Lung Cancer, Appl. Sci. 2019.

[4]. Kulkarni, A.; Panditrao, A. Classification of lung cancer stages on CT scan images using image processing. In Proceedings of the IEEE International Conference on Advanced Communications, Control and Computing Technologies, Ramanathapuram, India, 2014.

[5] J. R. Treadwell, H. M. Zafar, M. D. Mitchell, K. Tipton, U. Teitelbaum, and J. Jue, "Imaging tests for the diagnosis and staging of pancreatic adenocarcinoma: A meta-analysis," Pancreas, vol. 45, no. 6, pp. 789–795, Jul. 2016.

[6] Z. Q. Wang, J. S. Li, G. M. Lu, X. H. Zhang, Z. Q. Chen, and K. Meng, "Correlation of CT enhancement, tumor angiogenesis and pathologic grad ing of pancreatic carcinoma," World J. Gastroenterology, vol. 9, no. 9, pp. 2100–2104, Sep. 2003.

[7] M. S. Bhutani, P. Koduru, V. Joshi, P. Saxena, R. Suzuki, A. Irisawa, and K. Yamao, "The role of endoscopic ultrasound in pancreatic cancer screening," Endoscopic Ultrasound, vol. 5, no. 1, pp. 8–16, Jan./Feb. 2016.

[8] A. Dallongeville, L. Corno, S. Silvera, I. Boulay-Coletta, and M. Zins, "Initial diagnosis and staging of pancreatic cancer including main differentials," Seminars Ultrasound, CT MRI, vol. 40, no. 6, pp. 436–468, Dec. 2019. VOLUME 8, 2020 141715 M. Li et al.: Computer-Aided Diagnosis and Staging of PC Based on CT Images

[9] R. L. Siegel, K. D. Miller, and A. G. Sauer, S. A. Fedewa, L. F. Butterly, J. C. Anderson, A. Cercek, R. A. Smith, and and A. Jemal, "Cancer statis tics, 2020," Ca-a Cancer J. Clinicians, vol. 70, no. 1, pp. 7–30, Jan. 2020.

[10] K. Holub and C. Conill, "Unveiling the mechanisms of immune evasion in pancreatic cancer: May it be a systemic inflammation responsible for dismal survival?" Clin. Transl. Oncol., vol. 22, no. 1, pp. 81–90, Jan. 2020.

[11] M. Ilic and I. Ilic, "Epidemiology of pancreatic cancer," World J. Gas troenterology, vol. 22, no. 44, pp. 9694–9705, Nov. 2016.

[12] Q. J. Lin, F. Yang, C. Jin, and D. L. Fu, "Current status and progress of pancreatic cancer in China," World J. Gastroenterology, vol. 21, no. 26, pp. 7988–8003, Jul. 2015.

[13]. Huseyin Polat and Homay Danaei Mehr," Classification of Pulmonary CT Images by Using Hybrid 3D-Deep Convolutional Neural Network Architecture", Department of Computer Engineering, Faculty of Technology, Gazi University, 06500 Ankara, Turkey; h.danai.mehr@gmail.com , Published: 6 March 2019.

[14]. Nuthanakanti Bhaskar, Dr. Ganashree , "A Model: Lung Nodule Detection and Classification by SVM Network "Research Scholar, CSE Dept., VTU-RRC, Visvesvaraya Technological University, Belagavi, India, European Journal of Molecular & Clinical Medicine ISSN 2515-8260 Volume 7, Issue 8, 2020.

[15]. Diego Riquelme Moulay A. Akhloufi , "Deep Learning for Lung Cancer Nodules Detection and Classification in CT Scans ", Department of Computer Science, Université de Moncton, Moncton, NB E1A3E9, Canada; AI review, MDPI, 3 January 2020; Published: 8 January 2020.

[16]Society, AC: Cancer Facts & Figures. American Cancer Society (2016).

[17]   Sadot, E., Basturk, O., Klimstra, D.S., Gonen, M., Anna, L., Do, RKG, ¨DAngelica, M.I., DeMatteo, R.P., Kingham, T.P., Jarnagin, W.R., et al.: Tumor-associated neutrophils and malignant progression in intraductal papillary mucinous neoplasms: an opportunity for identification of highrisk disease. Annals of surgery 262(6), 1102 (2015).

[18] El-Baz, A., Nitzken, M., Khalifa, F., Elnakib, A., Gimelfarb, G., Falk, R., El-Ghar, M.A.: 3D shape analysis for early diagnosis of malignant lung nodules. In: IPMI. pp. 772–783. Springer (2011).

[19] Han, F., Wang, H., Zhang, G., Han, H., Song, B., Li, L., Moore, W., Lu, H., Zhao, H., Liang, Z.: Texture feature analysis for computer- aided diagnosis on pulmonary nodules. Journal of Digital Imaging 28(1), 99– 115 (2015)0278-0062 (c) 2018 IEEE.

[20] Way, T.W., Hadjiiski, L.M., Sahiner, B., Chan, H.P., Cascade, P.N., Kazerooni, E.A., Bogot, N., Zhou, C.: Computer-aided diagnosis of pulmonary nodules on CT scans: segmentation and classification using 3D active contours. Medical Physics 33(7), 2323–2337(2006).

[21] Lee, M., Boroczky, L., Sungur-Stasik, K., Cann, A., Borczuk, A., Kawut, S., Powell, C.: Computer-aided diagnosis of pulmonary nodules using a two-step approach for feature selection and classifier ensemble construction. Artificial Intelligence in Medicine 50(1), 43– 53 (2010).

[22] Kumar, D., Wong, A., Clausi, DA: Lung nodule classification using deep features in CT images. In: Computer and Robot Vision (CRV), 2015 12th Conference on. pp. 133–138. IEEE (2015).

[23] Buty, M., Xu, Z., Gao, M., Bagci, U., Wu, A., Mollura, DJ: Characterization of Lung Nodule Malignancy Using Hybrid Shape and Appearance Features. In: MICCAI. pp. 662–670. Springer (2016)

[24] Furuya, K., Murayama, S., Soeda, H., Murakami, J., Ichinose, Y., Yauuchi, H., Katsuda, Y., Koga, M., Masuda, K.: New classification of small pulmonary nodules by margin characteristics on highresolution CT. Acta Radiologica 40(5), 496–504 (1999).

[25] Recht, B., Fazel, M., Parrilo, P.A.: Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. SIAM review 52(3), 471–501 (2010).

[26] Zhou, J., Chen, J., Ye, J.: MALSAR: Multi-task learning via structural regularization (2012)

[27] Shalev-Shwartz, S., Tewari, A.: Stochastic methods for l1-regularized loss minimization. Journal of Machine Learning Research 12(Jun), 1865–1892 (2011).

[28] Nesterov, Y.: Introductory lectures on convex optimization: A basic course, vol. 87. Springer Science & Business Media (2013). [29]. B. Y. Rubinstein, Simulation and the Monte Carlo Method. John Wiley & Sons, New York, 1981.

[30]. I. Guyon, J. Weston, S. Barnhill, and V. Vapnik, "Gene selection for cancer classification using support vector machines," Machine Learning, vol. 46, pp. 389–422, 2002.

[31]. R. Gilad-Bachrach, A. Navot, and N. Tishby, "Margin based feature selection: theory and algorithms," in Proceedings of the 21st International Conference on Machine learning, 2004.

[32] Armato III, S., McLennan, G., Bidaut, L., McNitt-Gray, M.F., Meyer, C.R., Reeves, A.P., Zhao, B., Aberle, D.R., Henschke, C.I., Hoffman, E.A., et al.: The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): a completed reference database of lung nodules on CT scans. Medical Physics , 2011.

[33] Tran, D., Bourdev, L., Fergus, R., Torresani, L., Paluri, M.: Learning spatiotemporal features with 3D convolutional networks. In: ICCV. IEEE, 2015.

[34] Karpathy, A., Toderici, G., Shetty, S., Leung, T., Sukthankar, R., Fei-Fei, L.: Large-scale video classification with convolutional neural networks. In: IEEE CVPR, 2014.

[35] Tustison, N.J., Avants, B.B., Cook, P.A., Zheng, Y., Egan, A., Yushke vich, P.A., Gee, J.C.: N4ITK: Improved N3 bias correction. IEEE Transactions on Medical Imaging, 2010.

[36] Oliva, A., Torralba, A.: Modeling the shape of the scene: A holistic representation of the spatial envelope. IJCV, 2001.

Author Profile

Dr. G. Baskar, received his Master "s degree in Information Technology in KS.Rangasamy College of Technology, Tiruchengode, Tamil Nadu, India in 2008 and M.Phil Degree in Computer Science from Bharathiar University, Coimbatore, Tamil Nadu, India in 2010 and PhD degree in Computer Science from Government Arts College, Coimbatore, Tamil Nadu, India in 2016. His area of interest includes Data Mining, Image Processing and Bio Informatics.

R. Suba Lakshmi, is a PhD Research Scholar in Computer Science at KSG College of Arts & Science, Coimbatore and Assistant Professor,PSG College of Arts & Science Tamilnadu, India and received the bachelors, masters and mphil from Bharathiar University, Coimbatore, Tamil Nadu, India in 1999, 2008, 2013. Her research interest includes Data mining machine learning and support vector.