Intelligent Prediction of Lung Cancer via MRI/CT Images Using Morphological Neural Network Analysis

Vijayalakshmi.N[1], Ilayaraja.V[2],

M.E. Computer Science and Engineering, Department of Computer Science and Engineering, Shanmuganathan Engineering College, Thirumayam Taluk, Arasampatti, Pudukkottai, TamilNadu, India.

Assistant Professor, Department of Computer Science and Engineering, Shanmuganathan Engineering College, Thirumayam Taluk, Arasampatti, Pudukkottai, TamilNadu, India.

ABSTRACT: The main objective of this project is to assist clinicians to efficiently identify the Lung Cancer via forecasting analysis by using Morphological Neural Network (MNN) Classification with Image Pruning Methodology. In image processing domain, the complex task to analyze and rectify is cancer estimation and prediction. Compare to tumor estimations, Cancer estimations are more complex because these are purely cell based paradigms and visually not clear to analyze; especially this system focus on Lung cancer and its strategies. A new methodology is required to classify these kind of cancer cells, Morphological Neural Network Classifier with Image Pruning Scheme is introduced to efficiently classify the cancer cells and mark out the affected region more efficiently. Lung cancer is one of prevalent diseases in human, and can be diagnosed using several tests that include CT scans, MRI Scans, biopsy and so on. Over the years, the use of learning machine and artificial intelligence techniques has transformed the process of diagnosing lung cancer. However, the accurate classification of cancer cells is still a medical challenge faced by researchers. Difficulties are routinely encountered in the search for sets of features that provide adequate distinctiveness required for classifying breast tissues into groups of normal and abnormal. Therefore, the aim of this approach is to prove that the MNN algorithm is more efficient for diagnosis, prognosis and prediction of lung abnormality, which is basically derived from two classical algorithms called Morphological Image Processing and Artificial Neural Network (ANN) with Image Pruning.

KEYWORDS: Lung Cancer Prediction, Morphological neural Network, Artificial Neural Network, ANN.

I. INTRODUCTION

Due to large prevalence of smoking and air pollution around the world, lung cancer has become one of the most common and deadly disease in recent decades. It often takes long time to develop and most people are diagnosed with the disease within the age bracket 55 to 65. Early identification and treatment is the best available option for the infected people. Reliable identification and classification of lung cancer requires pathological test, namely, needle biposy specimen and analysis by experienced pathologists. However, because it involves human judgment of several factors and a combination of experiences, a decision support system is desirable in this case.
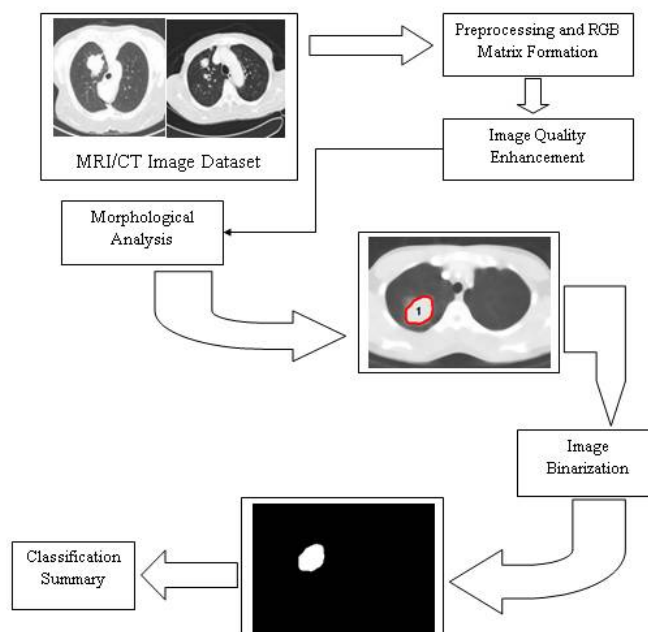
Fig.1 Proposed System Architecture

Recent developments in image processing, pattern recognition, dimensionality reduction and classification methods has paved the way for alternate identification and classification approaches for lung cancer. A number of machine learning (ML) algorithms including artificial neural network (ANN), support vector machine (SVM), discriminant analysis, decision trees and ensemble method addresses the diagnosis and classification of lung cancers through image processing and pathological identifiers. In addition to these ML approaches, deep learning through restricted Boltzmann machine in the form of autoencoders has shown promising success in classification tasks in different domain including acoustics, sentiment classification, and image and text recognition. Motivated by the success of deep learning in relevant fields, a deep learning based classification method is investigated in this work. The contribution of this work is two folds. Firstly, it is shown for the first time, deep learning methods can outperform existing works on small dataset for lung cancer classification. Secondly, an architecture of deep auto-encoder network for lung cancer classification is proposed which outperforms other methods and also show that the performance improvement is statistically significant.

II. LUNG CANCER PERCEPTIONS

As a result, more early stage lung cancers (i.e., stage I NSCLC) are detected during the regular lung cancer screening or other incident findings. Although early cancer detection and treatment can improve the survival rate of lung cancer patients, lung cancer recurrent rates after surgical resection of the malignant tumors can range from 30% to 60% as reported in the previous studies. As a result, mortality rate among the stage I NSCLC patients is also much higher than many other types of cancers (e.g., breast and colon cancer) detected at an early stage. According to the statistical data reported from the National Cancer Institute's Surveillance, Epidemiology, and End Results database, current 5-year survival rates are 49% and 45% for Stage IA and Stage IB NSCLC patients, respectively. Therefore, in order to more effectively treat and manage the stage I NSCLC patients, it is important to develop an effective clinical marker or prediction model to more accurately predict cancer prognosis after cancer surgery.

For the patients identified with the higher risk of cancer recurrence, the specific chemotherapy should be applied after surgery to minimize cancer recurrence risk. Since there is no clinical standard for assessing the risk of

post-surgery cancer recurrence to date, researchers have explored different genomic biomarkers to identify the molecular fingerprints of different genomic defects in lung cancer development and prognosis. Among them, two biomarkers namely, excision repair cross-complementing 1 (ERCC1) gene product and the regulatory subunit of ribonucleotide reductase (RRM1), have been extensively investigated and reported as prognostic biomarkers of the NSCLC patients. ERCC1 is a rate limiting enzyme in the nucleotide excision repair and RRM1 functions as a regulatory subunit of ribonucleotide reductase that controls the substrate specificity and the on/off function of ribonucleotide reductase.

III. CLINICAL PRACTICE

In the clinical practice, these two biomarkers are examined and measured by applying a standard Immune-histo-chemical (IHC) method to analyze the surgery-resected tumor tissue specimen. Studies have shown that high expression of ERCC1 together with RRM1 often indicated an improved outcome in NSCLC patients who are treated only with surgical resection. Thus, the researchers believe that comparative analyses between gene and IHC protein expression of ERCC1 and RRM1 might define a useful genomic biomarker to predict prognosis of the NSCLC patients. Despite the promising results, using genomic biomarkers still faces multiple challenges and has a number of limitations, such as the higher cost, processing error in signal probing and/or subjective rating, as well as the lower specificity. Hence, radiographic imaging examinations still play a critical role in current clinical practice for lung cancer detection, diagnosis, and prognosis assessment.

Many advanced imaging modalities including high-resolution CT, Positron Emission Tomography (PET), PET-CT and Magnetic Resonance Imaging (MRI) have been investigated and used in lung cancer imaging. Among them, CT remains the most popular imaging modality due to its higher accuracy, wide accessibility and cost-effectiveness. However, reading and interpreting a large number of CT images of a lung cancer case is difficult for radiologists. Using Response Evaluation Criteria in Solid Tumors (RECIST) guideline to subjectively measure tumor size in one-dimension and evaluate tumor size change during sequential CT examinations is often not reliable (due to the inter-reader variability) and has lower correlation to the clinical outcomes of the patients.

Therefore, there is an increasing interest in identifying novel quantitative image markers computed from the CT images to predict cancer prognosis. In this approach, we investigated a new quantitative image analysis method to predict the risk of lung cancer recurrence of the Stage I NSCLC patients after lung cancer surgery. For this purpose, we developed a new computer-aided detection (CAD) scheme to automatically segment malignant tumors depicting on CT images acquired before surgery and compute tumor-related morphological, density and texture based image features. Using a set of the selected non-redundant image features, we trained a Naïve Bayesian network based classifier to predict the risk of cancer recurrence of the Stage I NSCLC patients after surgical treatment.

IV. EXISTING APPROACHES – A SUMMARY

In the past system researchers face lots of struggle in the development of a unified approach that can be applied to all types of medical images and applications such as lung cancer estimation and tumor analysis. Even, the selection of an appropriate technique for a particular kind of medical image is a difficult problem. Thus, there is no universal accepted method for medical image segmentation process. So, it remains a challenging problem in image processing industry and computer vision fields. The most popular method used to estimate tumor and cancer cells is Fuzzy C-means, but the predictions falls into certain prescribed level not for a global mean, so that approach is again a problem to developers. Therefore, researchers/developers decide to combine two different algorithms to work out the result for cancer cell estimation and predictions.

V. PROPOSED SYSTEM SUMMARY

The proposed approach comprises three main stages: Preprocessing, Feature Extraction and Classification. The preprocessing stage involves techniques that include noise removal, enhancement and segmentation steps of the MRI/CT scanned images. The feature extraction stage entails five steps: wavelet decomposition, wavelet coefficient

extraction, normalization, energy computation, and coefficient reduction. Lastly, the classification stage involves the combinatorial use of Morphological Image Processing and Artificial Neural Network, called ' Morphological Neural Network with Image Pruning ', which is used to classify lung tissue into normal and abnormal. Computer programs or software created based on the human intellect can be used to aid doctors in decision making without conferring with specialists directly. The software was not developed to substitute the specialist or doctor, but to aid in the diagnosis and prediction of patients condition from specific regulations or "experience". Patients with high-risk factors or symptoms or predisposed to specific diseases or illness, could be selected to see a specialist for more treatment. Utilizing the technology particularly Artificial Intelligence (AI) techniques in medical applications could lower the cost, time involved, human proficiency and medical inaccuracies. For all our proposed approach aims to provide the complete solution to the medical image processing strategy with the help lung cancer estimation and prediction scenarios.
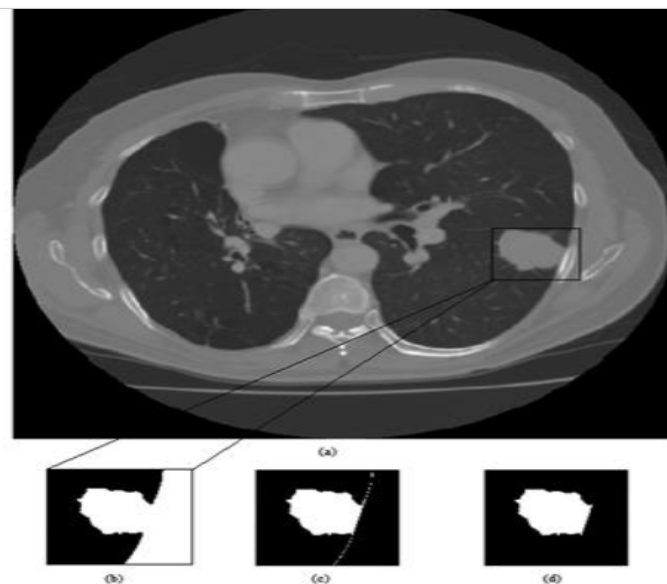
Fig.2 (a) A case with Juxtapleural tumor, (b) Leakage of region growing, (c) Chest wall connection is removed by convex hull function, (d) Final segmented region

VI. SYSTEM IMPLEMENTATION

The proposed system is designed with lots of purposeful modules that are described in detail:

A. Image Acquisition

The main purpose of this image acquisition module is to acquire the image and convert as Matrix format. The first stage of any vision system is the image acquisition stage. After the image has been obtained, various methods of processing can be applied to the image to perform the many different vision tasks required today. However, if the image has not been acquired satisfactorily then the intended tasks may not be achievable, even with the aid of some form of image enhancement. Image acquisition in image processing can be broadly defined as the action of retrieving an image from some source, usually a hardware-based source, so it can be passed through whatever processes need to occur afterward. Performing image acquisition in image processing is always the first step in the workflow sequence because, without an image, no processing is possible. The image that is acquired is completely unprocessed and is the result of whatever hardware was used to generate it, which can be very important in some fields to have a consistent baseline from which to work. One of the ultimate goals of this process is to have a source of input that operates within such controlled and measured guidelines that the same image can, if necessary, be nearly perfectly reproduced under the same conditions so anomalous factors are easier to locate and eliminate.

B. Image Preprocessing

The main necessity of this Image Preprocessing module is to enhance image quality. Image pre-processing can significantly increase the reliability of an optical inspection. Several filter operations which intensify or reduce certain image details enable an easier or faster evaluation. The following are the examples of image pre-processing method Normalization, Edge filters, Soft focus, selective focus, User-specific filter, Static/dynamic Binarization, Image plane separation and Binning. This modules comprises many internal works include: Includes several functions for image processing, Contrast increase by static or dynamic Binarization, look-up tables or image plane separation, Resolution reduction via binning, Image rotation and Conversion of color images to gray value images.

C. Image Binarization and Morphological Analysis

Binarization is the process of converting a pixel image to a binary image. The main contribution of this module is to Convert the gray image into binary and remove the unwanted contents of lung part. The images were firstly divided into normal and abnormal. Based on the fact that abnormal cases involve masses and microcalcifications, whether benign or malignant, 207 images were identified to be normal cases, while the remaining 115 images have abnormalities. The neural network was trained using 80% of both types, and tested with the remaining 20%. This procedure was performed four times repeatedly for each level of decomposition, from level 2 to level 5. For each level, features extracted using the reduction procedure were utilized. Binary images may contain numerous imperfections.

D. Image Pruning and ROI Selection and Segmentation

The pruning algorithm is a technique used in digital image processing based on mathematical morphology. It is used as a complement to the skeleton and thinning algorithms to remove unwanted parasitic components. In this case 'parasitic' components refer to branches of a line which are not key to the overall shape of the line and should be removed. These components can often be created by edge detection algorithms or digitisation. Lung image have two part (Left and Right) in which part we need to process of Left Lung and Right Lung. Image preprocessing techniques are essential for discovering the orientation of the mammogram, noise removal, and quality improvement of the image. Prior to the application of image processing algorithm on mammogram, preprocessing steps are very significant to focus the identification of abnormalities without unnecessary influence from background of the mammogram. For this study, the first phase of preprocessing entailed two procedures; noise elimination and image contrast enhancement. The second phase involves segmentation, which is executed to remove the background area (high intensity rectangular label, tape artifact, and noise). The third deals with the application of phase reduction and global gray level thresholding to extract region of interest (ROI).

E. Feature Extraction and Classification using Neural Network

The main motto of this feature extraction module is to extract the important data's from the image. This is the statistical feature and geometric like mean, brightness, skewness, kurtosis, area, perimeter. The function of feature extraction is to obtain new variables from the matrix of the image in order to create distinct classes. The feature extraction procedure involves the use of the wavelet decomposition process. These features are then classified in the next stage. Features extraction comprises five processing steps as illustrated as below, by which the features in the system are extracted from the wavelet coefficients. After getting these features we need to classify with trained data's. An artificial neuron network (ANN) is a computational model developed based on the configuration and functions of biological neural networks. Information transmitted through the network influences the structure of the ANN since a neural network evolves or learns based on the input and output. The field is referred to with several names that include connectionism, parallel distributed, processing, neuro-computing, natural intelligent system, and machine learning algorithm. ANN is an organized system of basic processing elements, nodes or units, whose functions are to some extent derived from the animal neuron. The processing ability of the network is controlled by the inter unit connection strengths, or weights, achieved via adaptation to, or learning from, a set of training patterns. The most essential elements of ANNs are modeled after the configuration of the human lungs.

VII. LITERATURE SURVEY

In the year of 2005, the authors "Singhal S1, Vachani A, Antin-Ozerkis D" proposed a paper titled "Prognostic implications of cell cycle, apoptosis, and angiogenesis biomarkers in non-small cell lung cancer: a review", in that they described such as: Lung cancer is the leading cause of cancer death in the U.S. with survival restricted to a subset of those patients able to undergo surgical resection.

However, even with surgery, recurrence rates range from 30% to 60%, depending on the pathologic stage. With the advent of partially effective, but potentially toxic adjuvant chemotherapy, it has become increasingly important to discover biomarkers that will identify those patients who have the highest likelihood of recurrence and who thus might benefit most from adjuvant chemotherapy. Hundreds of papers have appeared over the past several decades proposing a variety of molecular markers or proteins that may have prognostic significance in non-small cell lung cancer. This review analyzes the largest and most rigorous of these studies with the aim of compiling the most important prognostic markers in early stage non-small cell lung cancer. In this review, we focused on biomarkers primarily involved in one of three major pathways: cell cycle regulation, apoptosis, and angiogenesis. Although no single marker has yet been shown to be perfect in predicting patient outcome, a profile based on the best of these markers may prove useful in directing patient therapy.

In the year of 2009, the authors "R. Rosell and team" proposed a paper titled "Screening for epidermal growth factor receptor mutations in lung cancer", in that they described such as: Activating mutations in the epidermal growth factor receptor gene (EGFR) confer hypersensitivity to the tyrosine kinase inhibitors gefitinib and erlotinib in patients with advanced non-small-cell lung cancer.

We evaluated the feasibility of large-scale screening for EGFR mutations in such patients and analyzed the association between the mutations and the outcome of erlotinib treatment. From April 2005 through November 2008, lung cancers from 2105 patients in 129 institutions in Spain were screened for EGFR mutations. The analysis was performed in a central laboratory. Patients with tumors carrying EGFR mutations were eligible for erlotinib treatment.

In the year of 2013, the authors "L. Humphrey and team" proposed a paper titled "Screening for lung cancer with low-dose computed tomography: a systematic review to update the U.S. Preventive Services Task Force Recommendation", in that they described such as: Lung cancer is the leading cause of cancer-related death in the United States. However, persons with early lung cancer have lower lung cancer–related mortality than those with extensive disease, suggesting early detection and treatment of lung cancer might be beneficial. Low-dose computed tomography (LDCT) and chest x-ray (CXR) have been studied for early screening, with several new studies reporting results since the last review. To update the 2004 review of screening for lung cancer for the U.S. Preventive Services Task Force. MEDLINE (2000 to 2012), Cochrane Central Register of Controlled Trials and Cochrane Database of Systematic Reviews (through fourth quarter 2012), Scopus, and reference lists.

Details about participants, study design, analysis, followup, and results were abstracted; study quality was rated using established criteria, where applicable. Four trials reported the effectiveness of screening with LDCT for lung cancer in patients with personal smoking exposure: one large good-quality trial reported screening was associated with reduced lung cancer and all-cause mortality reductions of 20 percent (95% CI, 6.8 to 27.6) and 6.7 percent (95% CI, 1.2 to 13.6), respectively. Three small European trials (two fair- and one poor-quality) showed no benefit of screening. When the three good- or fair-quality trials were combined in random effects meta-analysis, the relative risk of lung cancer mortality was 0.81 (95% CI, 0.72 to 0.91). One trial evaluated CXR screening in over 150,000 participants from the general population and reported no benefit of screening in this group or in a subset with personal tobacco smoke exposure. The reported sensitivity of LDCT for detecting lung cancer ranged from 80 to 100 percent and specificity from 28 to 100 percent in six studies; each study varied in its reporting method.
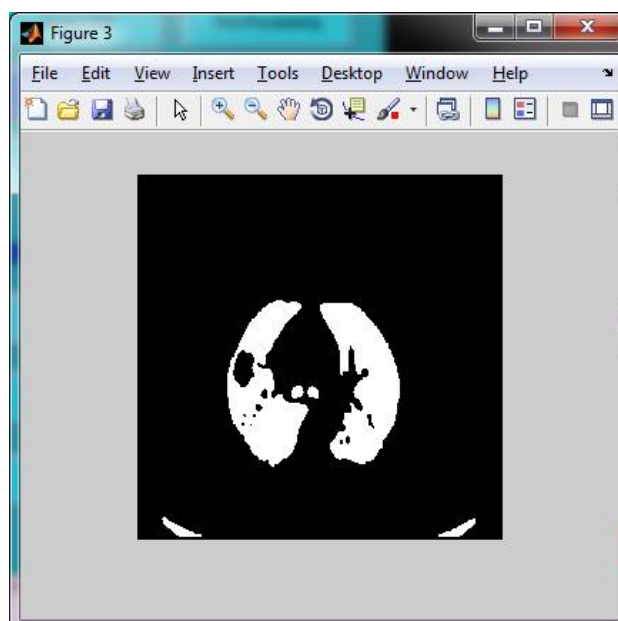
VIII. EXPERIMENTAL RESULTS

The following figure shows the Image Binarization concept of the proposed system.
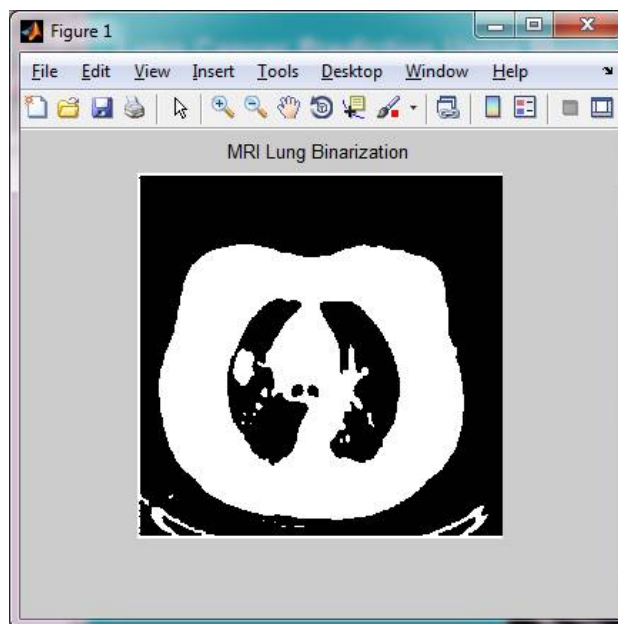
Fig.3 Image Binarization

The following figure illustrates the Image Segmentation Concept of the proposed system.
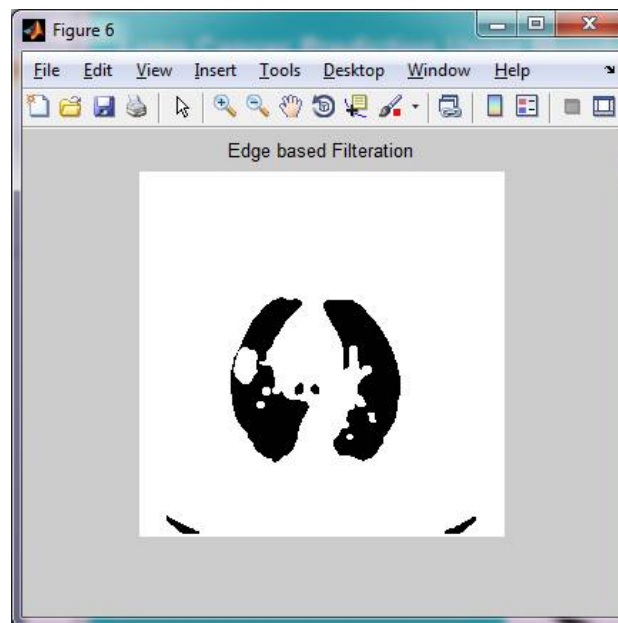

Fig.4 Image Segmentation

The following figure illustrates the filtering scenario of the proposed system.

Fig.5 Edge Filtering

IX. CONCLUSION

Developing precision medicine or a more effective personalized strategy for treating and managing Stage I NSCLC patients requires a more accurate clinical marker and/or assessment tool to predict cancer prognosis (including cancer recurrence risk). Current studies mainly focus on identifying more effective genomic biomarkers, demographic factors, and other clinical variables. In this study, we investigated a new quantitative image feature analysis approach using chest CT images and demonstrated two new study results. First, an image feature based classifier yielded significantly higher performance than two popular genomic biomarkers in predicting cancer recurrence risk. Second, the image features and genomic biomarkers are not highly correlated and provide supplementary information. As a result, fusion of these two types of features and biomarkers further improved prediction performance. Despite the limitations, this preliminary study provides us a valid foundation to continue working in this new and promising CAD field to develop and optimize highly performed and robust risk prediction schemes that may have potential to eventually assist clinicians in more accurately identifying the patients with a higher risk of lung cancer recurrence after surgery. Therefore, for these high risk patients, the post-surgery chemotherapy is required to prevent or minimize the risk of cancer recurrence and thus increase their disease-free survival and overall survival time.

REFERENCES

[1] Y. Qiang, Y. Guo, X. Li, Q. Wang, H. Chen, and D. Cuic, "The diagnostic rules of peripheral lung cancer preliminary study based on data mining technique," Journal of Nanjing medical university, vol. 21, no. 3, pp. 190–195, 2007.
[2] Z.-H. Zhou, Y. Jiang, Y.-B. Yang, and S.-F. Chen, "Lung cancer cell identification based on artificial neural network ensembles," Artificial Intelligence in Medicine, vol. 24, no. 1, pp. 25–36, 2002.
[3] J.-S. Lin, S.-C. B. Lo, A. Hasegawa, M. T. Freedman, and S. K. Mun, "Reduction of false positives in lung nodule detection using a two-level neural classification," Medical Imaging, IEEE Transactions on, vol. 15, no. 2, pp. 206–217, 1996.
[4] Y. P. Chiou, Y. F. Lure, P. Ligomenides et al., "Neural network image analysis and classification in hybrid lung nodule detection (hlnd) system," in Neural Networks for Processing [1993] III. Proceedings of the 1993 IEEE-SP Workshop. IEEE, 1993, pp. 517–526.

[5] K. Polat and S. G¨unes¸, "Principles component analysis, fuzzy weighting pre-processing and artificial immune recognition system based diagnostic system for diagnosis of lung cancer," Expert Systems with Applications, vol. 34, no. 1, pp. 214–221, 2008.

[6] M. G. Penedo, M. J. Carreira, A. Mosquera, and D. Cabello, "Computer aided diagnosis: a neural-network-based approach to lung nodule detection," Medical Imaging, IEEE Transactions on, vol. 17, no. 6, pp. 872–880, 1998.

[7] K. Mori, J.-i. Hasegawa, J.-i. Toriwaki, H. Anno, and K. Katada, "Recognition of bronchus in three-dimensional x-ray ct images with applications to virtualized bronchoscopy system," in Pattern Recognition, 1996., Proceedings of the 13th International Conference on, vol. 3. IEEE, 1996, pp. 528–532.

[8] K. Kanazawa, M. Kubo, N. Niki, H. Satoh, H. Ohmatsu, K. Eguchi, and N. Moriyama, "Computer aided screening system for lung cancer based on helical ct images," in Visualization in Biomedical Computing. Springer, 1996, pp. 223–228.

[9] "Computer assisted lung cancer diagnosis based on helical images," in Image Analysis Applications and Computer Graphics. Springer, 1995, pp. 323–330.

[10] R. Hayashibe, N. Asano, H. Hirohata, K. Okumura, S. Kondo, S. Handa, M. Takizawa, S. Sone, and S. Oshita, "An automatic lung cancer detection from x-ray images obtained through yearly serial mass survey," in Image Processing, 1996. Proceedings., International Conference on, vol. 1. IEEE, 1996, pp. 343–346.

[11] A. Khatami, A. Khosravi, C. P. Lim, and S. Nahavandi, "A wavelet deep belief network-based classifier for medical images," in International Conference on Neural Information Processing. Springer, 2016, pp. 467–474.

[12] A. Kavousi-Fard, A. Khosravi, and S. Nahavandi, "A new fuzzy based combined prediction interval for wind power forecasting," IEEE Transactions on Power Systems, vol. 31, no. 1, pp. 18–26, 2016.

[13] H. Zhou, H. Kong, L. Wei, D. Creighton, and S. Nahavandi, "Efficient road detection and tracking for unmanned aerial vehicle," IEEE Transactions on Intelligent Transportation Systems, vol. 16, no. 1, pp. 297–309, 2015.

[14] A. Khosravi and S. Nahavandi, "Load forecasting using interval type-2 fuzzy logic systems: Optimal type reduction," IEEE Transactions on Industrial Informatics, vol. 10, no. 2, pp. 1055–1063, 2014.

[15] A. Khosravi, S. Nahavandi, and D. Creighton, "Prediction interval construction and optimization for adaptive neurofuzzy inference systems," IEEE Transactions on Fuzzy Systems, vol. 19, no. 5, pp. 983–988, 2011.

[16] Z.-Q. Hong and J.-Y. Yang, "Optimal discriminant plane for a small number of samples and design method of classifier on the plane," pattern recognition, vol. 24, no. 4, pp. 317–324, 1991.

[17] L. Yu and H. Liu, "Feature selection for high-dimensional data: A fast correlation-based filter solution," in ICML, vol. 3, 2003, pp. 856–863.

[18] F. Badjio and F. Poulet, "Dimension reduction for visual data mining," in International symposium on applied stochastic models and data analysis (ASMDA-2005), 2005.