

Soft Computing Based Tumor and Lymph Node Detection in Thoracic Images

Tharani.V

Abstract –Lung is an important organ in our body which perform its function in both respiratory system and circulatory system. In recent years, Lung cancer is one of the most growing diseases in the world due to smoking. Early detection of Lung cancer leads to cure. But, early detection of Lung tumor through Medical Imaging techniques is not that much simple process. In this tumor detection, lymph nodes which are anatomical part of Human Lung cause false positives. In order to avoid false positives, a fully automatic differentiating method for Lung tumor and diseased lymph node from Computed Tomographic image of thoracic region. The detection and differentiation are performed in three stages, first detect all potential abnormalities in thoracic image, and then abnormalities are differentiated into lung tumor and diseased lymph nodes. Finally the lung tumor abnormalities are classified into Benign and Malignant tumors. These detection and differentiation are performed by Fuzzy logic and Neural Network to reduce false positive rate.

Index terms: Lung Cancer, Medical Imaging Techniques, Computer Tomography, Benign and Malignant tumors, Fuzzy logic, Neural Network.

1 INTRODUCTION

The lungs are the organs of respiration in humans. The main function of the lungs is to allow oxygen from the air to enter the bloodstream for delivery to the rest of the body. Through the lung, carbon dioxide is removed from the bloodstream and oxygen from inspired air enters the bloodstream. Lung cancer is caused by uncontrolled cell growth in tissues of the lung. If they left untreated, this growth can spread beyond the lung in a process called metastasis into nearby tissue and, eventually, into other parts of the body [4].

Lung cancer is the most common cause of cancer-related death in men and women, and is responsible for 1.3 million deaths annually, as of 2008[1]. In particular, nonsmall cell lung cancer (NSCLC) is the most prevalent type of lung cancer, accounting for about 80% of all cases. Staging, this assesses the degree of spread of the cancer from its original source, is the most important factor affecting the prognosis and potential treatment of lung cancer. For NSCLC, the tumor node metastasis (TNM) staging is the internationally agreed system, which involves analysis of the primary lung tumor, regional lymph nodes and distant metastases. The size and spatial extent of the primary lung tumor and the locations of the abnormal regional lymph nodes indicate a stage IA to IIIB NSCLC, while any distant metastases suggest a stage IV NSCLC.

Image processing is one of most growing research area these days. In the following an effective scheme to detect abnormal formation of cells in the lungs are detected and abnormality is diagnosed. This is an automatic approach that detects the tumor from lung Computed Tomographic (CT) image. In this work, objective is to design a fully automatic methodology for simultaneous detection of primary lung tumors and disease in regional lymph nodes from CT thoracic images.

The amount of publications on simultaneous detection of lung tumors and disease in regional lymph nodes is limited [1]. In our recent work, a region-based approach with spatial information was reported, which however, proposed a detection method mainly to facilitate image retrievals, rather than focusing on optimization of the detection performance.

Furthermore, the method required a separate class of tumor border, to work around the issue that the surrounding areas of tumors were often confused with the mediastinum [12]. Such a tumor-border class complicated the training process, which was quite unnatural for the clinical process. Later work avoided such an issue with a multilevel discriminative model and more comprehensive spatial features, but also posed several improvement opportunities. A similar type of work is on lung tumor detection, which first detects all abnormalities, then extracts only those that are highly representative of lung tumors [11]. By first segmenting the lung field, a threshold and fuzzy-logic based approach is then used to detect the lung tumors, but the detection performance is quite sensitive to the delineation accuracy of the lung field. Another approach attempts to handle tumors lying close to the edge of lung fields by incorporating the location, intensity, and shape information, but the method could potentially result in a large number of false positives with the predefined thresholds. To reduce the false positives detected in the mediastinum, learning-based techniques with tumor-specific features were proposed, but the methods were based on empirical studies of predefined threshold value and tumor sizes, and did not seem to consider abnormal lymph nodes in the thorax.

There are also a number of existing works on lymph node detection, mostly on CT images. Most of these methods utilize the segmentation of the anatomical structures in mediastinum, such as airways, aorta and pulmonary artery [8]. A Hessian matrix for detecting the blob-like shaped lymph nodes is also used. A deformable registration approach has been recently proposed to restrict the search area of the blob detectors, using a

Tharani.V is pursuing Master Degree in Applied Electronics in Anna University, Chennai, India, PH-9715257618 and E-mail: tharanibme48@gmail.com

probabilistic mediastinal lymph node atlas created by combining all database images with manually delineated lymph nodes. A different discriminative method was also proposed to detect lymph nodes based on comprehensive appearance and spatial features [6]. Detection methods for other types of lymph nodes include the directional difference filtering for abdominal nodes and the marginal space learning for axillary nodes. The detection performances of these approaches are usually highly related to the segmentation accuracy of anatomical structures, which is however, hard to avoid for CT images. These approaches also focus on the lymph nodes only, not considering cases with tumors, especially if they affect the appearances of the anatomical structures in the mediastinum. Another often studied area for lung tumor and regional lymph nodes is segmentation, including a number of different methods for tumor volume delineation on PET images with a comprehensive review in, those for CT, and PET-CT images, and lymph node segmentations on CT images.

An artificial neural network ensemble is a learning paradigm where several artificial neural networks are jointly used to solve a problem [10]. An artificial neural network ensemble to identify lung cancer cells in the CT lung images. First the Neural Network is trained with number of normal CT lung images and Abnormal CT lung images to differentiate normal and abnormal cells in the lung region. Then the Neural Network is trained to differentiate the abnormal cells into Benign and Malignant tumor cells. Benign are abnormal cells which does not leads to spreading of cancer cells to some other parts of the body, whereas Malignant are abnormal cells which leads to cancer cells spreading to neighbor organs.

Thus propose a multistage discriminative model to detect tumors and abnormal lymph nodes. First, we made no distinction between the two types, and classified the voxels into the lung field and abnormal region of interest with Neural Network Classifiers. Next, features in 3-D space to differentiate the detected abnormal volumes into tumors and abnormal lymph nodes. Last, it is formulated to refine the detected tumors for false positive reductions.

1.1 Outline

The series of operations involved in this work are Pre-processing, Segmentation, Feature Extraction, Tumor and Lymph node differentiation using Neural Networks and finally Tumor Classification.

Image pre-processing techniques are necessary, in order to find the orientation of the lung tumor, to remove the noise and to enhance the quality of the image. It refers to the process of partitioning digital image into multiple segments or sets of pixels. The goal of segmentation is to simplify or change the representation of an image into something that is more meaningful and easier to analyze. Feature extraction is defined as locating those pixels in an image that have some distinctive characteristics.

2. SUMMARY OF EXISTING SYSTEM

Designed an automatic methodology used for simultaneous detection of primary lung tumors and disease in regional lymph nodes from PET-CT thoracic images. The problem exhibits two main challenges. First, although PET indicates areas with high uptake activities, it can also highlight non-pathological areas (e.g., in myocardium), and the standard uptake value (SUV), which is a semi-quantitative measure of normalized radioactivity concentration, normally exhibits high inter-patient variances. Second, separations between lung tumors and abnormal lymph nodes are difficult. Although they may be differentiated by segmenting the lung fields from CT images, if tumors extent to the surrounding organs especially the mediastinum, such segmentations may not be reliable. For complex cases involving tumors invasion into the mediastinum or lymph nodes abutting the lung field, the ability to differentiate between the two types of abnormalities are more challenging.

The existing system using Conditional Random Fields (CRF) and Support Vector Machine (SVM) for detecting tumor in thoracic region of PET/CT image, which is an approximate detection since they are based on probability function. The PET/CT image is obtained by a combination of PET and CT imaging technique which undergoes the radiography with injection of radiation emitting tracer i.e.,F-fluoro-deoxy-glucose (FDG). Even though tracer helps in improving image quality, it has side effects in patients. The existing system has also less classification rate between normal and abnormal cells of lung region as well as benign and malignant tumor classification. This tends to have low accuracy rate. The procedure for this system takes high elapsed time for processing in SVM and CRF.

3. PROPOSED SYSTEM

The proposed method is an automatic method of detecting both tumors and abnormal lymph nodes simultaneously from Computed Tomography (CT) thoracic images. The detection is a multistage approach, by first detecting all potential abnormalities, then differentiates between tumors and lymph nodes, and finally refines the detected tumors for false positive reduction. Neural network based classification of lung tissues as benign or malignant.

The detection of tumor in existing system is based on uptake values of FDG by the normal and abnormal cells. This variation in uptake of tracer, FDG causes color variation in PET/CT image. This method of detection of tumor will not give more accurate result and since following steps of differentiating tumors also result in high false positive rate.

The proposed system uses CT thoracic image for detecting tumor which does not include tracer injection. This is the major advantage of this project as PET imaging

procedure is eliminated. Thus time taken for detection of tumors is less and use of Fuzzy rules based Neural Network classification of tumors achieves better classification rate with high performance. Also accuracy rate of the proposed system result when compared to existing system is high with low false rate.

4. SYSTEM IMPLEMENTATION

The system model starts with a typical Lung CT image which undergoes various image processing operations for image enhancement and feature extraction. In first step, the image is preprocessed to remove artifacts which refer

to noise in CT image. The preprocessed CT image is then endured with feature extraction. The features extracted are used for differentiation tumor region and lymph nodes in lung tissues. Finally Fuzzy based Neural Network is bring into play to perform classification of Normal and Abnormal lung tissues, it is further distinguished into Benign and Malignant tumors.

The proposed system uses Fuzzy logic and Neural Network for classification of lung tissues as benign or malignant. It achieves better classification rate with high performance. Preprocessing is the initial step for detecting the lung image.

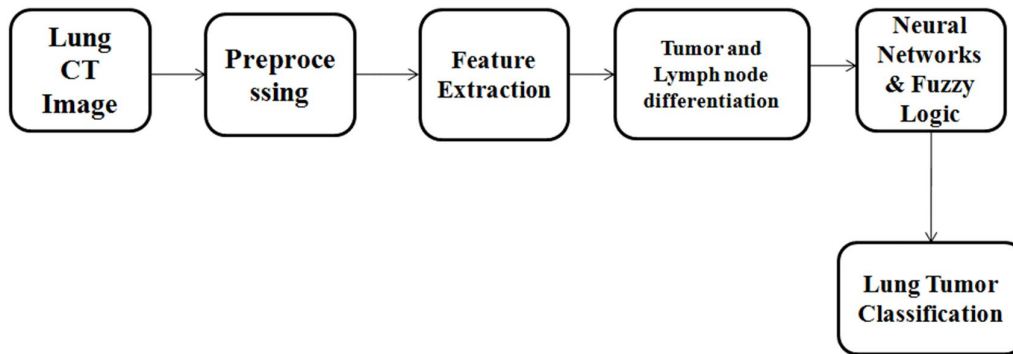


Figure1. Block diagram representing the Proposed System Model

The system implementation involves following steps:

- Preprocessing.
- Segmentation.
- Feature Extraction.
- Tumor and Lymph node differentiation.
- Lung tumor Classification using Fuzzy logic and Neural Networks Classifier.

5. PRE-PROCESSING

Image pre-processing techniques are necessary, in order to find the orientation of the lung tumor, to remove the noise and to enhance the quality of the image. Before any image-processing algorithm can be applied on CT thoracic image, preprocessing steps are very important in order to limit the search for abnormalities without undue influence from background of the image. CT thoracic images are medical images that are difficult to be interpreted, thus a preparation phase is needed in order to improve the image quality and make the segmentation results more accurate. The main objective of this process is to improve the quality of the image to make it ready to further processing by removing the unrelated and surplus parts in the back ground of the image.

White noise is one of the most common problems in image processing. Even a high resolution image is bound to have some noise in it. For a higher solution image a simple box blur may be sufficient, because even a tiny feature will be represented by a large group of pixels.

Pre-processing of CT lung image is the first step to reduce the noise and to enhance the image for further processing. The purpose of these steps is basically to improve the image and the image quality to get more surety and ease in segmenting the Lung Region alone. After enhancing the Lung CT image, the next step is to segment the Lung tumor region from Lung CT image.

Image pre-processing is the term for operations on images at the lowest level of abstraction. These operations do not increase image information content but they decrease it if entropy is an information measure. The aim of pre-processing is an improvement of the image data that suppresses undesired distortions or enhances some image features relevant for further processing and analysis task. Image pre-processing used to remove the redundancy in images. Neighbouring pixels corresponding to one real object have the same or similar brightness value. If a distorted pixel can be picked out from the image, it can be restarted as an average value of neighbouring pixels. Image pre-processing methods can be classified into categories according to the size of the pixel neighborhood that is used for the calculation of new

pixel brightness. In this paper, it will be presented some pixel brightness transformations and local pre-processing methods realized in MatLab.

5.1 Median filter

Due to low quality, low ability of distinguishing abnormalities from their surrounding, and artifacts which can be degrading the quality of CT images, sometimes to the point of making them diagnostically unusable; the first step in preprocessing is the image de-noising. Therefore the current step in this work is to perform application of median filter to the CT lung image.

The median filter is a nonlinear digital filtering technique, often used to remove noise. Such noise reduction is a typical pre-processing step to improve edge detection on an image. Median filtering is very widely used in digital image processing under certain conditions.

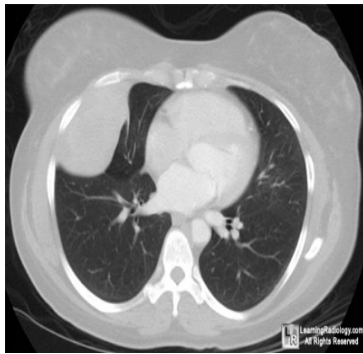


Figure 2. Lung CT original image

The median filter is a nonlinear digital filtering technique, often used to remove noise. Such noise reduction is a typical pre-processing step to improve the results of later processing (for example, edge on an image). Median filtering is very widely used in digital image processing because, under certain conditions, it preserves edges while removing noise.



Figure 3. Median filtered lung CT image

Many medical images may contain low-contrast, fuzzy contours. The histogram modes corresponding to the different types of regions in an image may often overlap and, therefore, segmentation by thresholding becomes difficult. Image pre-processing techniques can sometimes

help to improve the shape of the image histogram, by making it more strongly bimodal. The median filter replaces the value of each pixel by the average of all pixel values in a local neighborhood. In the median filter, the value of each pixel is replaced by the median value calculated in a local neighborhood.

5.2 Segmentation

It refers to the process of partitioning digital image into multiple segments or sets of pixels. The goal of segmentation is to simplify or change the representation of an image into something that is more meaningful and easier to analyze. Image segmentation is typically used to locate objects and boundaries in images. More precisely, image segmentation is the process of assigning a label to every pixel in an image such that pixels with the same label share certain visual characteristics. The result of image segmentation is a set of segments that collectively cover the entire image, or a set of contours extracted from the image. Each of the pixels in a region is similar with respect to some characteristic or computed property, such as color intensity, or texture. Adjacent regions are significantly different with respect to the same characteristics.

Segmentation is done to separate the image in to two or more sub module regions. Segmenting an image also saves the processing time for further operations which has to be applied to the image. This approach use segmentation using a global threshold in order to segment the tumor region from Lung CT image.

After enhancing the Lung CT image, the next step of this technique is to segment the Lung tumor region from Lung CT image. Segmentation is done to separate the image in to two or more sub module regions. Segmenting an image also saves the processing time for further operations which has to be applied to the image. Segmentation is using a global threshold in order to segment the tumor region from Lung CT image.

5.3 Morphological Closing Operation

Morphological close operation is applied on the thresholded image to fill in holes and small gaps in the image. Reserve the block whose area is the biggest and set the others to zero using 8-connected neighbors. The binary lung mask is obtained using the above step. Extract the lung boundary by setting a pixel to 0 if its 4-connected neighbors are all 1's, thus leaving only boundary pixels. Multiply the original Lung CT image with the lung masked image to obtain the final segmented lung region with gray level values as those of original image.

Mathematical morphology is used to process and analyze images. It process images based on shapes. It apply structuring element to an input image and produces the output image of an equivalent size.

5.4 Thresholding

Histogram thresholding is used for determination of the actual binary masks for the lung area. Binary masks are generated from input gray-level CT data using an iterative thresholding algorithm, a better method than the conventional thresholding algorithm, in which the threshold is simply chosen as the minimum between the two maxima of the gray level histogram.



Figure 4. Tumor Segmented Image by Thresholding

The image histogram is initially divided into two parts using a starting threshold value, which can be for example half the maximum of the dynamic range of the current image, or the conventional threshold value just described. Afterwards, the sample mean of the gray values associated with the foreground pixels and the sample mean of the gray values associated with the background pixels are computed, and a new threshold value is determined as the average of these two sample means. The process is repeated until the threshold value does not change any more.

6. FEATURE EXTRACTION

Feature extraction is defined as locating those pixels in an image that have some distinctive characteristics. Typically that characteristic is some inhomogeneity in local image properties. Step edges, for instance, are inhomogeneities in intensity or range.

The features extracted from images serve the following purposes:

- Comparison and similarity among images:
Quantitative comparison of anatomical structures and anomalies observed in medical imaging data is done. The comparison establishes relationships among observations by means of similarity measures.
- Classification of observations:
Observations such as modality, anatomical structure, or pathology are

classified to obtain more relevant search results, or to support steps such as training of structure specific vocabularies, and retrieval. The features are a basis for classifier learning, and application during retrieval.

- Identification and localization of anatomical structures:

Anatomical structures and groups of structures have to be identified both during retrieval, and during training, when unsupervised methods capture structure in the data set. The learning of this structure (e.g., similar anatomical regions across cases) is necessary to support efficient retrieval during application.

Features are used as inputs to classifiers that assign them to the class that they represent. In this feature extraction, Local Binary Pattern (LBP) features, Gray Level features, Wavelet features and Gray Level Co-Occurrence Matrix (GLCM) features are extracted.

7. LUNG TUMOR CLASSIFICATION USING FUZZY LOGIC AND NEURAL NETWORKS CLASSIFIER:

On lung cancer diagnosis, it is obvious that artificial neural networks have already been widely exploited in this area. To construct a classifier, a set of examples representing previous experience is essential. In general, the larger and more representative the set of available examples is, better classification of future query cases can be obtained. In the medical domain, however, there are several challenges and practical limitations associated with data collection. First, collecting data from patients is time-consuming. Second, acquiring large volumes of patients representing certain diseases is often challenging due to the low prevalence of the disease. Cancer prevalence is particularly low among screening populations which results in class imbalance in the collected set of examples; a phenomenon where one of the disease states is underrepresented. In addition, the clinical presentation of patients with the same disease varies dramatically. Classifier systems with neural networks are implemented replacing the rule base classifiers, in conventional classifier systems, by neural networks. The Classifier System is composed of a population of classifiers or rules. Associated to each classifier there is a "strength", used to express the energy or power of each classifier during the evolution process. Classifiers become unstable when trained by small size training dataset. The performance of most classifiers can be improved by increasing the size of the training dataset. Learning curves are commonly used to study the generalization properties of the trained classifiers as a function of the number of training data examples.

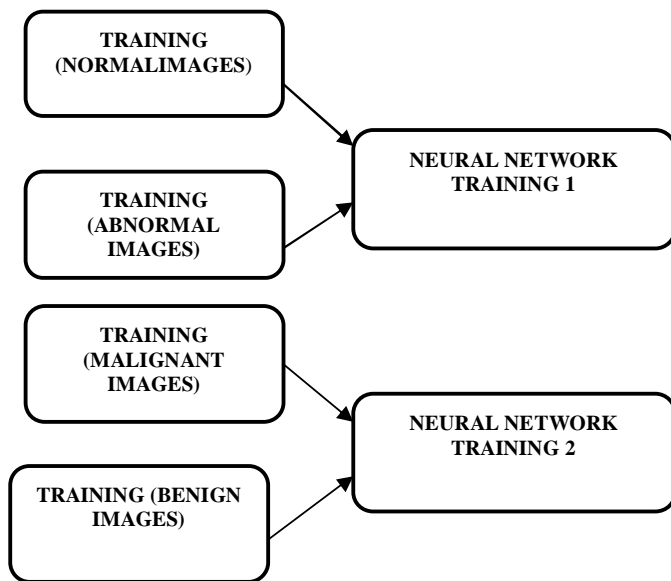


Figure 5. Training mode of Neural Network Classifier

The Neural Network Classifiers are trained to detect normal and abnormal cells in the lung region with the help of CT thoracic images. More number of CT thoracic images, both normal lung and diseased lung to train up the Neural Network Classifier. The detected abnormalities were differentiated as tumors or abnormal lymph nodes as shown in Figure 2. A Neural Network integrated with Fuzzy rules and a comprehensive set of features were designed to achieve an accurate discrimination between the two types of cells, normal lung tissues and abnormal diseased cells.

For that training of Neural Network number of CT thoracic images of various categories like normal lung image, abnormal lung image which has lung diseases like tuberculosis and abnormal lung image with lung cancer images are collected. The images are the datasets for the Neural Network training and finally the trained Neural network is used for differentiation of lymph nodes and lung tumor and as well as Lung tumor Classification.

8. CONCLUSION AND FUTURE WORK

A fully automatic detection and classification of lung tumor from CT thoracic images have presented. The Neural Network Classifiers are trained with more number of CT thoracic images to improve the accuracy rate of detection. The Fuzzy logic rules are used in training of the classifiers which make the classification between normal and abnormal as well as benign and malignant with less false

positive rate. The false negative rate is reduced up to the level of 9% and accuracy rate is improved with 20 normal and 40 abnormal such as 20 benign and 20 malignant tumors which is about 96%. The results that obtained from this Neural Network Classifier is accurate when compared to any other method. The processing time for detection and classification of lung tumors also less and when Neural Network Classifier trained with more number of CT images, then its classification is 2-5 times faster than SVM (Support Vector Machine) and CRF (Conditional Random Fields) methods. This kind of faster detection and classification have effective application in Lung tumor Diagnosis and Lung tumor Surgery.

This work can be improved by further reducing false negatives of abnormal lymph nodes which is misidentified as lung tumor. I have planned to extend this method of lung tumor detection and classification to lung MRI images and PET images. This extension can make extraction and lung tumor detection in 3-D thoracic images. Then CT of any other images like brain, abdomen and uterus can also be used to train the Neural Network for tumor diagnosis in these regions.

ACKNOWLEDGMENT

The author would like to thank Radiology Department in Bharat Scans in Chennai, India and Faculty member Dr. B. Santharaj, Radiological Safety Officer.

REFERENCES

- [1] Yang Song, Weidong Cai, Jinman Kim, and David Dagan Feng, 2011, " Multistage Discriminative Model for Tumor and Lymph Node Detection in Thoracic Images".
- [2] W. De Wever, S. Stroobants, J. Coolen and J.A. Verschakelen, 2009, "Integrated PET/CT in the staging of nonsmall cell lung cancer: technical aspects and clinical integration"
- [3] Cherry Ballangan, Xiuying Wang, Michael Fulham, Stefan Eberl, Yong Yin, and Dagan Feng, 2011, "Automated Delineation of Lung Tumors in PET Images Based on Monotonicity and a Tumor-Customized Criterion"
- [4] Poonam Bhayan, Gagandeep Jindal, 2011, "A Segmented Morphological Approach to Detect Tumor in Lung Images"
- [5] Manaswini Padhan, 2011, "An Extensive Survey on Artificial Neural Network Based Cancer Prediction Using Soft-Computing Approach"
- [6] Johannes Feulner, S. Kevin Zhou, Martin Huber, Joachim Hornegger, Dorin Comaniciu, Alexander Cavallaro, "Lymph Node Detection in 3-D Chest CT using a Spatial Prior Probability"
- [7] Yuri Boykov, and Vladimir Kolmogorov, 2004, "An Experimental Comparison of Min-Cut/Max-Flow Algorithms for Energy Minimization in Vision"
- [8] Jan-Martin Kuhnigk, Volker Dicken, Lars Bornemann, Annemarie Bakai, Dag Wormanns, Stefan Krass, and Heinz-Otto Peitgen, 2006,

- "Morphological Segmentation and Partial Volume Analysis for Volumetry of Solid Pulmonary Lesions in Thoracic CT Scans"
- [9] Zhi-Hua Zhou, Yuan Jiang, Yu-Bin Yang, Shi-Fu Chen, 2002, "Lung Cancer Cell Identification Based on Artificial Neural Network Ensembles"
- [10] K.A.G. Udeshani, R.G.N. Meegama, T.G.I. Fernando, "Statistical Feature-based Neural Network Approach for the Detection of Lung cancer in Chest X-Ray Images"
- [11] Yang Song, Weidong cai, Stefan Eberl, Michael J.Fulham and Dagan Feng, "Region and Learning based Retrieval for Multi-Modality Medical Images"
- [12] Hanford J. Deglinc, Rangaraj M. Rangayyan, Fábio J. Ayres, Graham S. Boag, Marcelo K.Zuffo, 2007, "Three-Dimensional Segmentation of the Tumor in Computed Tomographic Images of Neuroblastoma"
- [13] Li Zhang, Weida Zhou, and Licheng Jiao, 2004, "Wavelet Support Vector Machine"
- [14] S. Limmer, C. Stöcker, V. Dicken, S. Kraß, H. Wolken and P. Kujath "Computer-Assisted Visualization of Central Lung Tumours Based on 3-Dimensional Reconstruction"
- [15] W. Wever, S. Stroobants, J. Coolen, and J. Verschakelen, 2009, "Integrated PET/CT in the staging of nonsmall cell lung cancer: Technical aspects and clinical integration," Eur. Respir. J., vol. 3.
- [16] Y. Song, W. Cai, S. Eberl, M. Fulham, and D. Feng, 2011 "Discriminative pathological context detection in thoracic images based on multi-level inference," vol. 6893.
- [17] C. Cortes and V. Vapnik, 1995, "Support-vector networks," Mach. Learn., vol. 20.
- [18] J. Lafferty, A. McCallum, and F. Pereira, "Conditional random fields: Probabilistic models for segmenting and labeling sequence data"
- [19] Y. Song, W. Cai, S. Eberl, M. Fulham, and D. Feng, 2011 "Thoracic image case retrieval with spatial and contextual information"
- [20] I. Jafar, H. Ying, A. Shields, and O. Muzik, 2006, "Computerized detection of lung tumors in PET/CT images"
- [21] Y. Cui, B. Zhao, T. Akhurst, J. Yan, and L. Schwartz, 2008, "CT-guided automated Detection of lung tumors on PET images"
- [22] C. Ballangan, X.Wang, S. Eberl, M. Fulham, and D. Feng, 2010, "Automated lung tumor segmentation for whole body PET volume based on novel downhill region growing"