

# Lung Cancer Diagnosis from CT Images Based on Local Energy Based Shape Histogram (LESH) Feature Extration and Pre-processing



Denny Dominic and K. Balachandran

**Abstract** Lung cancer as of now is one of the dreaded diseases and it is destroying humanity never before. The mechanism of detecting the lung cancer will bring the level down of mortality and increase the life expectancy accuracy 13% from the detected cancer diagnosis from 24% of all cancer deaths. Although various methods are adopted to find the cancer, still there is a scope for improvement and the CT images are still preferred to find if there is any cancer in the body. The medical images are always a better one to find with the cancer in the human body. The proposed idea is, how we can improve the quality of the diagnosis form using pre-processing methods and Local energy shape histogram to improve the quality of the images. The deep learning methods are imported to find the varied results from the training process and finally to analyse the result. Medical examination is always part of our research and this result is always verified by the technicians. Major pre-processing techniques are used in this research work and they are discussed in this paper. The LESH technique is used to get better result in this research work and we will discuss how the image manipulation can be done to achieve better results from the CT images through various image processing methods. The construction of the proposed method will include smoothing of the images with median filters, enhancement of the image and finally segmentation of the images with LESH techniques.

**Keywords** LESH techniques · Pre-processing · Feature extraction · Histogram · Segmentation

---

D. Dominic (✉) · K. Balachandran  
Department of Computer Science and Engineering, Christ (Deemed to be University),  
Bengaluru, India  
e-mail: [denny.cmi@gmail.com](mailto:denny.cmi@gmail.com)

K. Balachandran  
e-mail: [balachandran.k@christuniversity.in](mailto:balachandran.k@christuniversity.in)

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

A. Kumar and S. Mozar (eds.), *ICCCE 2020*, Lecture Notes in Electrical Engineering 698, [https://doi.org/10.1007/978-981-15-7961-5\\_68](https://doi.org/10.1007/978-981-15-7961-5_68)

## 1 Introduction

Due to poor prognostic methods the diagnosis of the lung cancer has become difficult. The survival is possible only if the detection could make it early [1]. Cancer can be spread through the blood vessels, or lymph fluid and they are surrounded by cancer tissues around the lungs. There is always the possibility of the growth of cancer cells around the chest form the lung nodules. The major two categories of lung cancer are Non-small cell lung cancer and small cell lung cancer.

The four stages of the cancer cells vary by their way of shape and size cancer cells. Smoking is found to be the reason for lung cancer and they are best found though the CT images. Thus, it is necessary to have a best frame work to find out the cancer from the CT images.

The system should be able to pre-process the image and they should be able to make an effort to categories them according to the size and the features that are attributed to it. Local energy map is developed form the histogram that is generated by LESH techniques.

Pre-processing is a must because the images are collected from the different machines and they may differ in their format. Therefore, it is necessary to the desired result will require the pre-processing on the images. The purpose of the paper is to develop an energy map, based on the LESH technique and they should be involved with pre-processing methods with latest pre-processing techniques. The medical images are always with different layers' inner structure and they need to process them with different characteristic to get desired result [1]. Pre-processing is a must, to get the intended result and they shall be of the metrics values, binary numbers, pre-processed images and charts.

The chest computed tomography (CT) scan, has high resolution thus, it is used for lung related diseases. It is difficult to find the small cell malignancy at the early stages. There is always an urgent task to develop a frame work for it because they have to be formulated and they have to be identified. The major areas where the pulmonary nodules develops are difficult to find out because there are in the inner most part of the body and thus the location of the malignancy. The pulmonary nodules are identified based on the shape. Hence, it is an important task to identify the various CT images from various sources.

The important task of this research is that, to know different pre-processing method and developing feature extraction method with the support of the LESH techniques.

## 2 Literature Survey

LESH technique has been very successful in applying multi-disciplinary fields related to pattern recognition and as part of the lung cancer detection to differentiate many features they are ought to detect [2-4]. The noise in the image is always a

matter of disturbance and they can be separated as in two steps and they are Weiner filtering and Denoising. In the image processing white noise is a problem [5]. IlyaLevner [6] has observed that the topographical and other noises can be reduced using watershed segmentation.

The pixel grouping and watershed segmentations are always used in the image processing and computer vision related research. When the object of the pixels with same proximity are in the same vicinity the image mapping is necessary and this can be done using energy mapping with LESH techniques. Non overlapping region can be used with the help of the natural phenomena and they can be done with segmentation methods. The idea about the quantification for semantic nodule characteristics of sphericity, speculation and clarification have become major concern [7–9]. Hence, the image quantification has to be done with image segmentation. Anywhere in the body the nodule can occur. In the image processing, the idea if there is any occurrence of abnormality has to be found out according to the useful information [10] and it becomes the foundation of image classification.

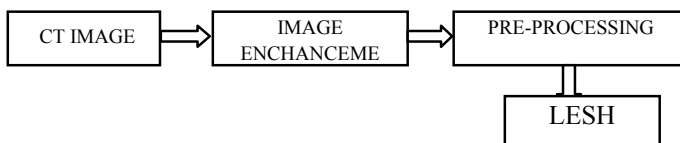
LESH feature extractions for the similarity of abnormal nodules in the CT images have achieved high classification accuracy. The experiments have done with wavelet and local energy based shape histogram (LESH) and the main features have amalgamated with classifiers and finally the confusion matrix is generated [11].

GINNEKEN [10] has shown a different approach in executing the rule based and pixel classification approaches. The preferred approach is rule based by most of the research category [12–15], with different steps, the rule can be used for the extraction of the result. LESH features are mainly used for extracting the images and it is calculating the CT image by the histogram from the local energy pattern.

### 3 Methodology

The images are pre-processed with the support of LESH vector with an aided supported of image extraction, enhancement and finally segmentation techniques. The extractions of the features are done with morphological operation and colorimetric operations (Fig. 1).

The Structure of Image Enhancement is the system which enhances the image using median filter, Gabor filter and watershed segmentation. The enhancement of colour and clarity can be increased by these methods. Nucleuses are achieved by the thresholding methods in the images to segment. At the end of all, the above given



**Fig. 1** Structure of Image Enhancement

process, the system will achieve colorimetric and morphological to image of the nucleuses is extracted feature [5]. The extracted morphologic features use average intensity, area, perimeter and eccentricity of the nucleuses. The reasons to select the CT images are, because they are with more accuracy and less distortion. Images are acquires form various health care centres.

### 3.1 Structure of Image Enhancement

#### 3.1.1 Median Filter

The median filter will allow the high value spatial frequency and it is proved that it will remove the noise from the images. There is always half of the pixels are being affected by the noise removal and this process may result in wrong information about the raw data. Median filter also used for blurring on edges and this process can be applied to the image to enhance the methods that are useful in the process.

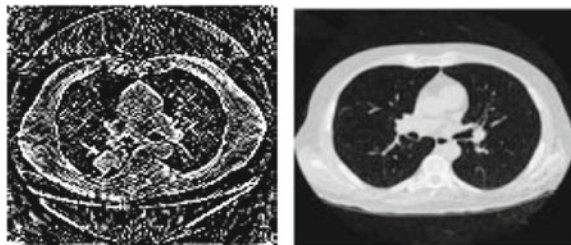
Median filtering in a non-linear operation and removes the salt and pepper results from the images. The result contains median value in the neighbourhood around the corresponding pixel in the image. The 0's at the end will pad on the images as shown in Fig. 2.

#### 3.1.2 Gabor Filter

Gabor filter is used to have wider perception about the input image thus it can interpret to final results. Although the spatial and frequency domains still exist for the enhancement but we prefer Gabor filter which can give better results due to its wider perception capacity. Gabor filter has better performed in our research than auto enhancement and Fast Fourier [2].

2D Gabor filters are constructed using the following factors: the wavelength  $\lambda$  of the sinusoidal function and the orientation  $\theta$  of the normal to the parallel stripes of the Gabor function 4. The spatial aspect ratio  $\gamma$  specifies the ellipticity of the support of the Gabor function.

Fig. 2 Zero Component Analysis (ZCA) Whitening



ZCA\_Whitening

Original Image

The variance  $\sigma$  of the Gaussian function. The equation for 2D representation of Gabor filter is given as the expansion of the equation,  $x' = x\cos\theta + y\sin\theta$ ;  $y' = -x\sin\theta + y\cos\theta$ , For  $\gamma = 1$ , circular is the support. For  $\gamma < 1$ , the orientation of the parallel stripes of the function is being supported elongated.

### 3.1.3 Watershed Segmentation

In the watershed segmentation images are segmented into multiple images for better understanding. Thus watershed segmentation is segmented to objects and constituent regions. Thus the segmentation finally segments images into multiple images and as a set of contours extracted from the image [5]. Marker controller is used for watershed segmentation and they convert the segmented images into white or grey images for that matter. In this segmentation approach the frontier and regional methods are combined. The working of the segmentation is explained as the RGB images are converted to grey scale images and then they are processed for edge detection and structural elements and a foreground markers erosion is developed with background markers. Here the regional maxima are developed for watershed segmentation and dilation is done on the final image to get visualization of the images. The Fig. 3 shows the results of watershed segmentation in the Confusion Matrix.

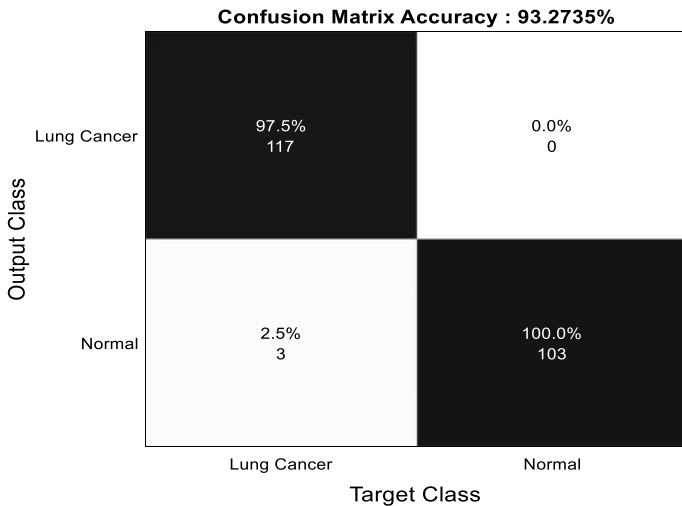


Fig. 3 Confusion Matrix

### 3.1.4 Feature Extraction Based on LESH

Based on calculating of local energy pattern of the histogram for the image of interest. Local energies are calculated along with different orientations.

Before calculate phase congruency (PC), the Gabor filter is involved with different orientations  $o$  and scales  $s$ . Transfer function for log-Gabor is given

$$G(\omega) = \frac{(\log(\frac{\omega}{\omega_0}))^2}{2(\log(\frac{k}{\omega_0}))^2} \quad (1)$$

where  $\omega_0$  is central frequency of the filter and  $\frac{k}{\omega_0}$  is a constant.

Let  $G_{so}^{\text{even}}$  and  $G_{so}^{\text{odd}}$  be the even-symmetric and odd-symmetric filters at orientation  $o$  and scale  $s$ . Then convolution with image result of response vector given as

$$[e_{so}(z), o_{so}(z)], \quad (2)$$

where  $z = (x, y)$  represents a location. Hence, the orientation can be computed with the amplitude of the response as

$$A_{so} = \sqrt{(e_{so}(z))^2 + (o_{so}(z))^2} \quad (3)$$

And sensitive phase deviation measure is given as:

$$\Delta\Phi_m(z) = \cos(\Phi_m(z) - \bar{\Phi}_m(z)) - |\sin(\Phi_m(z) - \bar{\Phi}_m(z))| \quad (4)$$

Now, local energy is calculated as:

$$E(z) = \sqrt{(\sum m)E_{so}(z))^2 + (\sum m)O_{so}(z))^2} \quad (5)$$

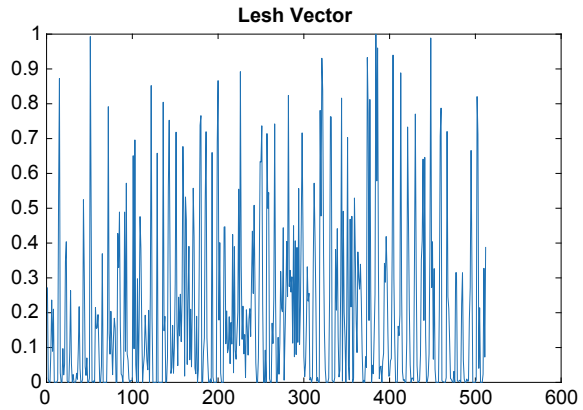
At last, 2D phase congruency for the image is normalized by sum of amplitude computed Fourier amplitude components as

$$PC(z) = \frac{E(z)}{\sum m A_{so} + \epsilon} \quad (6)$$

$$PC(z) = \frac{\sum m W(z)[A_m(z)\Delta\Phi_m(z) - T]}{\sum m A_m + \epsilon} \quad (7)$$

where T-noise cancellation factor and W(z)- weighting of the frequency spread.  $A_m$  and  $\phi_m$  represent the amplitude and phase angle respectively.

**Fig. 4** Lesh Vector



Further, the resultant of LESH feature vector is calculated as follows:

$$H_{r,b} = \sum W_r \times PC(z) \times \delta_{r-b} \tag{8}$$

$$W_1 = \frac{1}{\sqrt{2\Pi}\sigma} e^{[(x-rx_0)^2 + (y-ry_0)^2]/\sigma^2} \tag{9}$$

where  $W_r$ - Gaussian weighting function of the delta of the orientation. The histogram of Lesh vector is shows in Fig. 4.

## 4 Conclusion and Future Scope

This paper has discussed mainly about how the LESH vector can be applied to an image so as to get a better result. It is done based on the local energy patterns that have been developed through the histogram. Different orientations have been used for it to calculate the local energy pattern. The result of the research shows that the image is processed with 2-D log-Gabor filter for better orientation  $o$  and scales  $s$ . The image enhancement is done always with the idea in mind that to have better interpretability on the image. Enhancements are usually done on the spatial domain and frequency domain. Watershed segmentation and Gabor filters are always used for accurate results for pre-processing stage and hence to get a better result from the LESH vector. Thus the region of interest is used to find the area of cancer. For the further work the Convolutional neural network can give better result and it will increase the accuracy and detection of cancer nodules with high clarity.

## References

1. Cruz-Roa A, Caicedo JC, González FA (2011) Visual pattern mining in histology image collections using bag of features. *Artif Intell Med* 52(2):91–106
2. Sarfraz MS, Hellwich O (2008) Head pose estimation in face recognition across pose scenarios. In: International Conference on Computer Vision Theory and Applications VISAPP, vol 1, pp 235–242
3. Sarfraz MS, Hellwich O (2008) An efficient front-end facial pose estimation system for face recognition. *Int J Pattern Recog Image Anal* 18:434–441
4. Sarfraz MS, Hellwich O (2009) On head pose estimation in face recognition. In: *Computer Vision And Computer Graphics. Theory And Applications. Lecture Notes CCIS*, vol 24, pp 162–175
5. Chaudhary A, Singh SS (2012) Lung cancer detection on CT images using image processing. In: *International transaction On Computing Sciences*, vol. 4
6. Levner I, Zhang H (2007) Classification-driven watershed segmentation. *IEEE Trans Image Process* 16(5):1437–1445
7. Farag A, Ali A, Graham J, Elshazly S, Falk R (2011) Evaluation of geometric feature descriptors for detection and classification of lung nodules in low dose CT scans of the chest. In: 2011 IEEE International Symposium, Paper presented at the Biomedical Imaging: From Nano to Macro, Chicago, IL, USA, 30 March to 2 April 2011
8. Lin P-L, Huang P-W, Lee C-H, Wu M-T (2013) Automatic classification for solitary pulmonary nodule in CT image by fractal analysis based on fractional Brownian motion model. *Pattern Recog* 46(12):3279–3287
9. Farag A, Elhabian S, Graham J, Farag A, Falk R (2010) Toward precise pulmonary nodule descriptors for nodule type classification. In: *Medical Image Computing and Computer-Assisted Intervention (MICCAI 2010)*. Springer, Berlin
10. Ginneken BV, Romenyand BM, Viergever MA (2001) Computer-aided diagnosis in chest radiography: a survey. *IEEE Trans Med Imaging* 20(12):1228–1241
11. Wajid SK, Hussain A (2015) Local energy-based shape histogram feature extraction technique for breast cancer diagnosis. *Expert Syst with Appl* 42(20):6990–6999
12. Sharma D, Jindal G (2011) Identifying lung cancer using image processing techniques. In: *International Conference on Computational Techniques and Artificial Intelligence (ICCTAI 2011)*, vol 17, pp 872–880
13. Nguyen HT et al (2003) Watersnakes: energy-driven watershed segmentation. *IEEE Trans Pattern Anal Mach Intell* 25(3):330–342
14. Ansari MD, Ghrera SP (2018) Intuitionistic fuzzy local binary pattern for features extraction. *Int J Inf Commun Technol* 13(1):83–98
15. Ansari MD, Ghrera SP (2016) Feature extraction method for digital images based on intuitionistic fuzzy local binary pattern. In: 2016 International Conference System Modeling & Advancement in Research Trends (SMART). IEEE, pp 345–349 November 2016