CT Imaging Technology for Lung Cancer Detection: A Mechanism-Based Approach

Amanjot[1], Meenakshi Sharma[2]

[1]Research Scholar, Department of Computer Science and Engineering, RIEIT, Ropar
[2]Assistant Professor, Department of Computer Science and Engineering, RIEIT, Ropar

Abstract: ***Addressing image classification remains a significant challenge in image pre-processing methodologies. Numerous classification approaches have been developed to tackle this issue. This study introduces a novel classifier, SVM-FA (Support Vector Machine optimized with Firefly Approach), specifically designed for detecting lung cancer based on CT images. Given the critical nature of lung cancer, early analysis of the disease is imperative. To achieve this, the study employs image pre-processing techniques, including filtration and segmentation, on input CT scan images. Subsequently, the SVM classifier optimized with the firefly approach is applied to the pre-processed data with the goal of improving the accuracy of the final prediction. To assess the effectiveness of the proposed SVM-FA approach, a comparative analysis is conducted, comparing it with traditional methods and the standard SVM classifier. The results indicate that the proposed approach is both effective and efficient, achieving a high accuracy rate (98%) and specificity (84%) in comparison to the alternatives.***

Keywords: Pre-processing of medical images, Support Vector Machine (SVM) classification, Optimization using Firefly Algorithm, Lung cancer

1.Introduction

A tumor signifies an abnormal growth of new tissue that can occur in any organ of the human body. Various types of cancer, such as lung, bone, and brain cancer, are currently considered significant threats globally, with varying survival rates [1-2]. Lung cancer is characterized by uncontrollable cell replication leading to tumor development. Removal of cancer cells can occur through the alveoli in the hemoglobin or via the lymphatic system surrounding lung tissue. The lymph travels through lymphatic vessels to lymph nodes in both the alveoli and the middle of the chest. As the lymph flows from the lungs to the chest's center, the lung cancer tumor progresses towards the chest. Metastasis occurs when tumor cells escape their original location and travel to lymph nodes or other parts of the body via the bloodstream [2]. Initial cancerous growth into the alveoli is known as primary cancerous. Furthermore, this cancer is classified into two major stages: small cell and non-small cell lung cancer further subcategorized as Carcinoma, Adenocarcinoma, and Squamous cell carcinomas. In 2008, among Jordanians, the classification of cancers for both males and females revealed more than 300 cases, representing 7.7% of all newly diagnosed cancer cases in the same year. Specifically, 297 males (13.1%) and 59 females (2.5%) were diagnosed with lung cancer, resulting in a male-to-female ratio of 5: 1. Consequently, this cancer ranks second among males and tenth among females [3].

In general, lung cancer detection involves four stages. Firstly, normal and abnormal CT images are gathered from an accessible database on the IMBA web page [5]. In the second stage, image enhancement methodologies are applied to obtain high-quality, clear images. Subsequently, image segmentation algorithms are implemented, playing a significant role in image processing. In the final stage, general characteristics are extracted from the improved segmented image, indicating whether the images are normal or abnormal

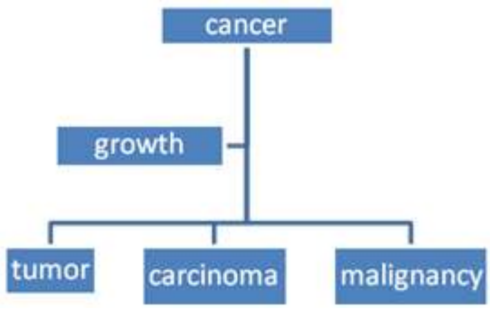


Figure 1: Lung Cancer block diagram

2.Related Work

Asuntha. A et al. [1] conducted an analysis involving edge detection, segmentation, and filtering to extract features from images. They subsequently developed a model to process these features, aiming to reduce obstacles and isolate the desired shapes within the images.

Mokhled Altarawneh [2] proposed a feature detection method to enhance the accuracy of image pixels. This emphasis on quality and accuracy is crucial for maintaining image stability in the research context.

Mokhled S. Al-Tarawneh [5] presented a model for cancer diagnosis employing Image Processing techniques. In this approach, Gabor filter was utilized for denoising the images, with notable effectiveness. Two segmentation methods, namely marker-controlled watershed segmentation and thresholding approach, were applied. Marker-controlled segmentation demonstrated superior results compared to the thresholding approach. Image feature extraction involved binarization, and a masking approach was employed to identify cancerous regions.

Anita Chaudhary [6] proposed a method for cancer detection through the processing of scanned images. Gabor filter was employed for noise reduction, and segmentation was achieved using marker-controlled watershed and threshold techniques. Tumor identification relied on extracting features such as area, perimeter, and roundness.

Nooshin Hadavi and Md. Jan Nordin [7] developed a diagnostic method for CT-Scan images based on Cellular Learning Automata (CLA). Gabor filter was employed for noise removal, and segmentation was carried out using the Region Growing algorithm. Features extracted from image segments were then applied to a new cellular automata algorithm for cancer identification.

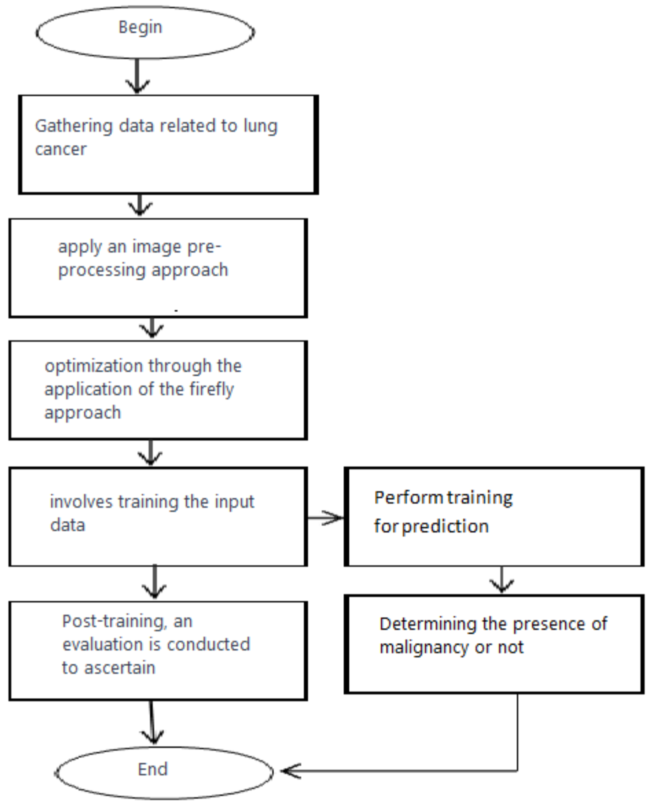
Aggarwal, T. et al. [8] proposed a model utilizing gray-level characteristics, optimum thresholding for segmentation, and extraction of geometrical features. Linear Discriminate Analysis (LDA) was employed as a classifier to differentiate between nodules and pulmonary structures. Despite employing simpler image processing techniques, the model achieved 84% accuracy, 53.33% specificity, and 97.14% sensitivity.

3.Proposed Work

Medical image processing stands out as a prominent research area, particularly in the medical field, where it is utilized for diagnosing diseases based on various imaging modalities such as MRI, CT scans, and X-rays. Before reaching a final diagnosis from these images, several pre-processing techniques are applied, including feature extraction and classification.
A review of the existing literature reveals that conventional methods for lung cancer detection using CT scans involve cancer diagnosis through the application of SVM Gabor filters, watershed segmentation techniques, and SVM classification approaches. While the traditional approach has proven to be effective, there is still room for improvement to enhance the accuracy of the results obtained. The limitation identified in traditional methods is the utilization of SVM classifiers for classification, necessitating the development of a novel classification approach.

To address this limitation, the current study employs image pre-processing techniques and an SVM classification approach. In the proposed work, optimization of SVM classifier factors is achieved using the firefly optimization approach. Improvements in SVM are introduced by altering the values of sigma and bounding box boundaries from 1. However, varying these factors is not a straightforward task. To facilitate these variations, the concept of the firefly optimization technique is implemented in this work. The flowchart of the proposed work is illustrated in Figure 2.

- Begin
- The initial step involves gathering data related to lung cancer, presented in the form of images. These images serve as inputs for the proposed SVM-FA approach.
- Upon obtaining the input images, the subsequent step is to apply an image pre-processing approach. This involves filtration and segmentation techniques, followed by cross-validation on the pre-processed images.
- In this phase, the SVM classification approach undergoes optimization through the application of the firefly approach.
- Following optimization, the next stage involves training the input data, focusing on predictive modeling.
- Post-training, an evaluation is conducted to ascertain whether malignancy is detected. In this context, lung cancer is identified in its malignant form.
- End



4.Image Classification and Optimization

A. Support Vector Machine (SVM)

Image classification stands as a fundamental concern within image processing. The objective of image classification is to predict the ordering of initial images based on their inherent characteristics. To address this challenge, various methodologies such as Adaptive Boost (AdBoosted), K-Nearest Neighbors (KNN), Artificial Neural Networks (ANN), and Support Vector Machines (SVM) can be employed.

SVM, in particular, holds significance in pattern image classifications. The SVM design is crafted to segregate a set of training images into distinct stages, represented as (x1, y1), (x2, y2), . . ., (xn, yn), where xi exists in Rd (d-dimensional feature space) and yi in {-1, +1} denotes the class label, with i ranging from 1 to n [1]. It constructs optimal hyperplanes based on a kernel function (K). Once the points are classified by the hyperplane, images on one side correspond to class-1, while those on the other side correspond to class +1 [9].

Paper ID: SR24110155303 DOI: https://dx.doi.org/10.21275/SR24110155303

B. Firefly Optimization

The Firefly algorithm involves two variables: the intensity of light and attractiveness. Fireflies are attracted to others with a brighter flash. Thus, both factors operate concurrently.

As a result, attractiveness and the specific distance (r) from the light source are inversely related. Consequently, an increase in distance leads to a decrease in both light and attractiveness.

The pseudo code for the Firefly algorithm is outlined as follows:

Objective Function for Accuracy Improvement, f(x):
Initialize the population of fireflies with parameters Sigma and Constraints.
Determine the fitness or intensity of light I_i for each firefly in the population based on f(x_i).
Define the absorption coefficient of light γ.
For each iteration (t=1 to maxIterations):
For each firefly i (i=1 to n, where n is the number of flies):
For every other firefly j (j=1 to n):
If (I_i < I_j), firefly i will move towards firefly j:
Vary attractiveness with distance r using exp[-γr].
Verify new solutions obtained and update the intensity of light for the fireflies.
End loop for firefly j.
End loop for firefly i.
Classify fireflies and determine the best solution g.
End loop for iterations.

5.Result Analysis

This study utilizes the firefly-optimized Support Vector Machine (SVM) approach for the detection of lung cancer from medical images. The outcomes obtained from the application of the proposed image processing mechanism for lung cancer detection are presented in this section. The performance of the proposed methodology is quantified in terms of accuracy, sensitivity, and specificity, as illustrated in the subsequent graphs. Additionally, a comparative analysis is conducted, comparing the traditional method, SVM approach, and the proposed SVM approach to demonstrate the effectiveness of the proposed work.

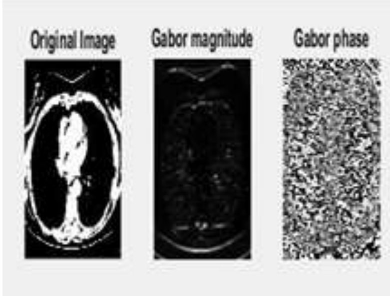


Figure 3a: Original Image with gabor filter

Figure 3b; Gradient Image

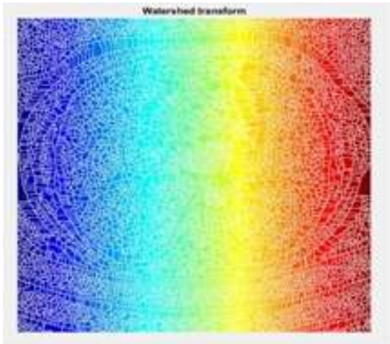


Figure 3c: Watershed Transform

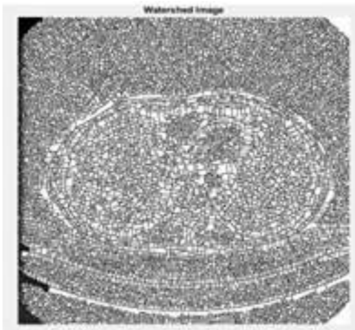


Figure 3d; Watershed Image

Figure 3e: Reconstruction Image

Figure 3f: Segmented Image

Figure 3: Various steps in image processing

Figure 5 illustrates a graphical comparative analysis between the traditional method, SVM, and the proposed SVM approach in terms of accuracy. The formulation for evaluating the accuracy of the proposed approach is presented below.

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \ldots\ldots (1)$$

Actual Values

Predicted Values		
	TP=19	**FP=0**
	FN=1	**TN=5**

6.Conclusion and Future Scope

In contemporary times, the utilization of images has become ubiquitous across various domains for detection and analysis purposes. Medical image processing, specifically, is a domain where diseases are diagnosed based on diverse images such as CT scans MRI, X-rays, etc. This research introduces a novel approach for detecting lung cancer through the utilization of CT scans. The optimization of the SVM classifier is achieved through the application of the firefly optimization approach. The proposed work's performance is evaluated based on efficiency, specificity, and sensitivity. Comparative analysis highlights the performance of the proposed approach against traditional classifiers and SVM classifiers concerning the considered performance metrics.

In the future, enhancements to the execution of the proposed work can take various forms, including efforts to reduce complexity. Additionally, there is potential for further improvements in feature selection and extraction mechanisms.

References

[1] Asuntha, A., Banu, P. A., Ainthaviarasi, K., Kumar, B. S., & Srinivasan, A. (2017). Feature extraction to detect Bone Cancer Using Image Processing. Research Journal of Pharmaceutical Biological and Chemical Sciences, 8 (3), 434-442.

[2] Non-Small Cell Lung Cancer, Available at: http://www. katemacintyrefoundation.org/pdf/non-small-cell.pdf, Adapted from National Cancer Institute (NCI) and Patients Livingith Cancer (PLWC), 2007, (accessed July 2011).

[3] Tarawneh, M., Nimri, O., Arqoub, K., & Zaghal, M. (2007). Cancer incidence in Jordan. Jordan National Cancer Registry.

[4] Lung Cancer Database, Available at: https://eddie. via. cornell. edu/cgibin/datac/signon. cgi, (accessed July 2011).

[5] Al-Tarawneh, M. S. (2012). Lung cancer detection using image processing techniques. Leonardo Electronic Journal of Practices and Technologies, 11 (21), 147-58.

[6] Chaudhary, A., & Singh, S. S. (2012, September). Lung cancer detection on CT images by using image processing. In 2012 International Conference on Computing Sciences (pp.142-146). IEEE.

[7] Hadavi, N., Nordin, M. J., & Shojaeipour, A. (2014, June). Lung cancer diagnosis using CT-scan images based on cellular learning automata. In 2014 International Conference on Computer and Information Sciences (ICCOINS) (pp.1-5). IEEE.

[8] Aggarwal, T., Furqan, A., & Kalra, K. (2015, August). Feature extraction and LDA based classification of lung nodules in chest CT scan images. In 2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI) (pp.1189-1193). IEEE.

[9] Thai, L. H., Hai, T. S., & Thuy, N. T. (2012). Image classification using support vector machine and artificial neural network. International Journal of Information Technology and Computer Science (IJITCS), 4 (5), 32-38.

[10] Makaju, S., Prasad, P. W. C., Alsadoon, A., Singh, A. K., & Elchouemi, A. (2018). Lung cancer detection using CT scan images. Procedia Computer Science, 125, 107-114