iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 7 **Issue:** IV **Month of publication:** April 2019

DOI: https://doi.org/10.22214/ijraset.2019.4664

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.177
Volume 7 Issue IV, Apr 2019- Available at www.ijraset.com

Comparison of GLCM and GLRLM for Lung Cancer Identification

Prof. Mrs. K. ChandraPrabha[1], S. Kausalya[2]
[1]*Associate Professor & Head of the Department, Department of Computer Science and Engineering, ACGCET, Karaikudi, Tamil Nadu.*
[2]*PG Student, Department of Computer Applications, ACGCET, Karaikudi, Tamil Nadu.*

Abstract: ***Lung cancer prevalence is one of the highest of cancers, at 18 %. One of the first steps in lung cancer diagnosis is sampling of lung tissues or biopsy. These tissue samples are then microscopically explored. This procedure is taken once imaging tests indicate the presence of cancer cells in the chest. Microscopic lung biopsy images are in RGB format which is converted into gray scale images. Gray scale images are examined for texture extraction using the Gray Level Co-Occurrence Matrix (GLCM) method used to obtain texture parameters of contrast, correlation, energy, and homogeneity features and Gray Level Run Length Matrix (GLRLM) method used to obtain parameters of SRE, GLN, RLN and RP features. Images are classified into two classes of cancer and non-cancer using Convolutional Neural Network (CNN) algorithm. This system compares the result of the accuracy of the Gray Level Co-occurrence Matrix (GLCM) and Gray Level Run Length Matrix (GLRLM) method.***
Keywords: Microscopic biopsy, GLCM, GLRLM, CNN, Classification.

I. INTRODUCTION

Lung malignant growth is one of the commonest tumors in the industrialized world, and people with this grave malady must arrangement with the physical impacts as well as with the psychosocial viewpoints. Lung malignant growth is an ailment of strange cells increasing and developing into a tumor. Among various sorts of malignant growth the lung disease is the most forceful and best practice to its exact anticipation is the assurance of the flow phase of the infection. A standout amongst the most vital and troublesome errands a specialist needs to do is the location and finding of harmful lung knobs from x-beam picture's outcome. Given that lung disease is one of the normal malignant growths around the world, the ramifications of concentrating on personal satisfaction just as survival require to be comprehended. Early location is the most essential for decreasing the demise because of lung malignant growth. The lung malignant growth conclusion is the example of lung tissues or biopsy. This technique can improve the precision and productivity for lung disease discovery. The point of this exploration is to plan a lung disease identification framework dependent on examination of tiny picture of biopsy utilizing advanced picture preparing. Minuscule lung biopsy pictures are in RGB group which is changed over into dim scale pictures. Dim scale pictures are broke down for surface extraction utilizing the Gray Level Co-Occurrence Matrix (GLCM) technique used to get surface parameters of difference, connection, vitality, and homogeneity highlights and Gray Level Run Length Matrix (GLRLM) strategy used to acquire parameters of SRE, GLN, RLN and RP highlights. Pictures are grouped into two classes of malignancy and non-disease utilizing Convolutional Neural Network (CNN) calculation. This framework looks at the consequence of the precision of the Gray Level Co-event Matrix (GLCM) and Gray Level Run Length Matrix (GLRLM) technique.

II. METHODOLOGY

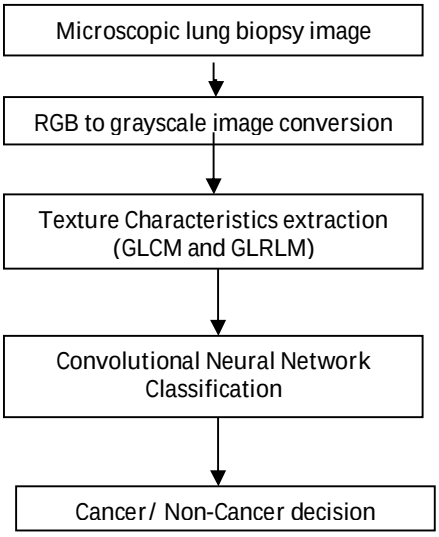


Fig. 1 Block Diagram of the Lung Cancer Identification System

A. Image Acquisition

Lung malignant growth is a standout amongst the most widely recognized and savage infections on the planet. Identification of lung disease in its beginning time is the key of its fix. When all is said in done, measures for beginning time lung malignant growth analysis for the most part incorporates those using X-beam chest films, CT, MRI, isotope, bronchoscopy, and so forth., among which an imperative measure is the supposed neurotic finding that investigations the examples of needle biopsies acquired from the groups of the subjects to be examined. The lung pictures are transferred to conclusion the lung malignant growth. Attractive Resonance Images utilized in the biomedical to distinguish and envision better subtleties in the inward structure of the body. Biomedical imaging and restorative picture handling that assumes an essential job for biopsy pictures has now turned into the most testing field in building and innovation. In this module, client can enter the MRI picture with different size and different sorts. Pictures are transferred as prepared and testing sets.

B. Conversion

Gray scale is the collection or the range of monochromic (gray) shades Dim scale is the accumulation or the scope of monochromic (dim) shades, going from unadulterated white on the lightest end to unadulterated dark on the contrary end. Dark scale just contains luminance(brightness) data and no shading data; that is the reason most extreme luminance is white and zero luminance is dark; everything in the middle of is a shade of dim. The dark scale pictures contain just shades of dim and no shading. Dim scale is otherwise called colorless at the most grounded. Dim scale is a scope of shades of dim without evident shading. The darkest conceivable shade is dark, which is the complete nonattendance of transmitted or reflected light. The lightest conceivable shade is white, the all out transmission or impression of light at all noticeable wavelengths.

C. Feature Extraction

The process of image features extraction is carried out with texture analysis using the Gray Level Co-Occurrence Matrix (GLCM) and Gray Level Run Length Matrix (GLRLM) method. This method works on the principle of calculating the probability of nearest neighbour between two pixels on certain distance and angular orientation. Co-occurrence means happening at the same time. Distance is stated as pixels, while orientation is in degrees. GLCM and GLRLM feature extraction is carried out in 4 angular direction, each with a 45º interval. They are; 0º, 45º, 90º, and 135º, whereas the distance between two pixels is given as1 pixel. The GLCM feature extraction method is a matrix that describes the occurrence frequency of two pixels with certain intensities at distance d and angular orientation θ within an image. The GLCM functions characterize the textures of an image by calculating how often a pair of the pixel with gray level or value i occur either horizontally, vertically, or diagonally to adjacent pixels with the value j (i and j represent the grey level values in the image). After creating the GLCM, several texture features are derived from the images like contrast, correlation, homogeneity and energy are calculated on the co-occurrence matrix.

TABLE I
FEATURE EXTRACTION

Features	**GLCM**				**GLRLM**			
Images	**Contrast**	**Correlation**	**Energy**	**Homogeneity**	**SRE**	**GLN**	**RLN**	**RP**
1	0.98	1.89	1.25	1.54	1.20	1.16	1.05	1.15
2	1.27	0.99	1.13	1.23	0.97	1.52	1.76	1.09
3	1.09	1.06	0.89	1.05	1.12	1.02	1.03	0.95
4	1.43	1.29	0.99	1.22	1.11	1.01	1.32	1.03
5	1.29	1.06	1.24	1.26	0.98	1.22	1.20	0.97

1) GLCM:

i) Contrast: Contrast features are used to calculate the degree of difference of grayness in an image. The greater the difference of grayness, the higher the contrast is. On the contrary, the less significant the difference of grayness between two pixels, the lower the contrast will be. Contrast is defined as:

$$\text{Contrast} = \sum_i \sum_j (i-j)^2 p(i,j) \quad (1)$$

where p(i, j) is the GLCM matrix

ii) Correlation: Correlation brings out how correlated a reference pixel to its neighbor over an image. Correlation is defined as:

$$\text{Correlation} = \sum_i \sum_j \frac{ijp_d(i,j)-\mu_x\mu_y}{\sigma_x\sigma_y} \quad (2)$$

Where μx , μ y and σ x ,σ y are the mean and standard deviations of probability matrix GLCM along row wise x and column wise y.

iii) Energy: Energy value describes the degree of grayness distribution in an image. Energy is written as:

$$\text{Energy} = \sum_i \sum_j p^2(i,j) \quad (3)$$

iv) Homogeneity: Homogeneity features calculate the degree of homogeneity of grayness in an image. Homogeneity is defined as:

$$\text{Homogeneity} = \sum_i \sum_j \frac{p(i,j)}{1+|i-j|} \quad (4)$$

2) GLRLM:

i) Short Run Emphasis : Short Runs Emphasis (SRE) fragments each run length value by the length of the run squared. This tends to emphasize short runs. The denominator is the overall number of runs in the image and provides as a normalizing factor.

$$\text{SRE} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \frac{P(i,j|\theta)}{i^2}}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} P(i,j|\theta)} \quad (5)$$

ii) Gray Level Non-uniformity: It computes the similarity of gray level values all the way through the image. The GLN is expected small if the run lengths are alike throughout the image.

$$\text{GLN} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} P(i,j|\theta)^2}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} P(i,j|\theta)} \quad (6)$$

iii) Run Length Non-uniformity: It computes the similarity of length of runs all the way through the image. The RLN is expected small if the run lengths are alike throughout the image.

$$\text{RLN} = \frac{\sum_{j=1}^{N_g} \sum_{i=1}^{N_r} P(i,j|\theta)^2}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} P(i,j|\theta)} \quad (7)$$

iv) Run Percentage : It computes the homogeneity and the distribution of runs of an image in a particular direction. RP is the largest when the length of runs is 1 for all gray levels in specific direction.

$$\text{RP} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \frac{P(i,j|\theta)}{N_p} \quad (8)$$

D. Classification

In cutting edge investigation of therapeutic imaging utilizing radiomics, machine, and profound picking up, including convolutional neural systems (CNNs), has been investigated. These methodologies offer incredible guarantee for future applications for both indicative and prescient purposes. CNNs are non-expressly modified calculations that recognize pertinent highlights on the pictures that enable them to group an info object. Connected in different assignments, for example, recognition (e.g., bosom sores on mammographic filters), division (e.g., liver and liver injuries on figured tomography (CT)), and determination (e.g., lung sores on screening low-portion CT). CNNs are an AI strategy dependent on a counterfeit neural system with profound design depending on convolution activities (the straight utilization of a channel or bit to nearby neighborhoods of pixel/voxels in an info picture) and down inspecting or pooling tasks (gathering of highlight map signals into a lower-goals include map). The last order or relapse task depends on more elevated amount highlights illustrative of an expansive open field that is smoothed into a solitary vector. The improvement of a calculation involves (a) determination of the hyper parameters, (b) preparing and approval, and (c) testing. The hyper parameters incorporate the system topology, the quantity of channels per layer, and the advancement parameters. Amid the preparation procedure, the dataset of info pictures (partitioned into preparing and approval sets) is over and over submitted to the system to catch the structure of the pictures that is striking for the undertaking. At that point, they are balanced at every cycle, focusing on minimization of the misfortune work, which evaluates how close the expectation is to the objective class. The execution of the prepared model is then assessed utilizing an autonomous test dataset. This is likewise gone for evaluating whether an "over fitting" has happened.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.177
Volume 7 Issue IV, Apr 2019- Available at www.ijraset.com

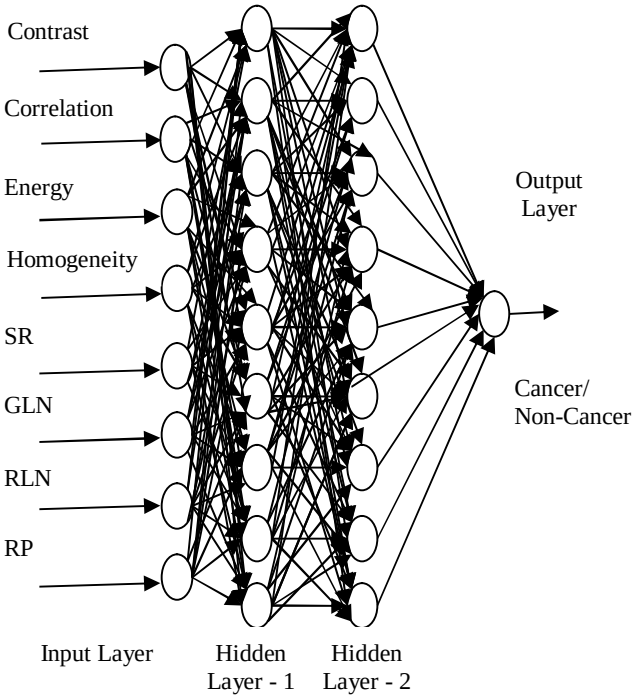


Fig. 2 Architecture of the Convolutional Neural Network

III. RESULT AND DISCUSSION

A. Microscope Lung Biopsy Image

The process of image features extraction is carried out with texture analysis using the Gray Level Co-Occurrence Matrix (GLCM) and Gray Level Run Length Matrix (GLRLM) method. These method works on the principle of calculating the probability of nearest neighbour between two pixels on certain distance and angular orientation and continuous pixels. This approach builds co-occurrence matrices of image data and GLRLM is the set of continuous pixels having same gray level, which in turn determine features as the matrix function of those images.

Some samples of microscopic lung biopsy images

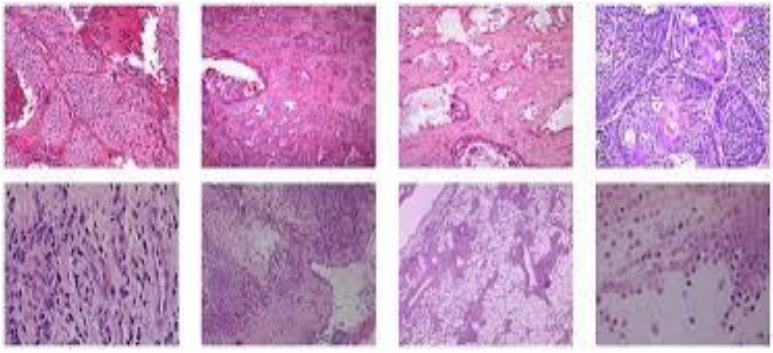


Fig. 3 Samples of Microscopic Lung Biopsy Image (Above: Cancer and Non-Cancer)

Co-occurrence means happening at the same time. This translates to the probability of one level of a pixel value being nearest to a value level of another pixel at certain distance (d) and angular orientation (θ) and the Run length is the number of neighbouring gray levels in particular direction. Distance is stated as pixels, while orientation is in degrees. Orientation is made up of four angular directions, each with a 45° interval. They are; 0°, 45°, 90°, and 135°, whereas the distance between two pixels is given as1 pixel.

A co-occurrence matrix is a square matrix whose number or elements is the square of pixel intensity level on an image. Then from that co-occurrence matrix, parameters of contrast, correlation, energy, and homogeneity are extracted as texture features and for GLRLM is a way of extracting higher order statistical features in the lung images. GLRLM is the set of continuous pixels having same gray level. The run length is the number of neighbouring gray levels in particular direction. The

features Short Run Emphasis (SRE), Gray Level Non-uniformity (GLN), Run length non-uniformity (RLN), Run percentage (RP) are used.

IV. ACCURACY

Results from both training and testing stages of the developed artificial neural network algorithm show that this algorithm is capable of properly classifying microscopic lung biopsy images into either cancer or non-cancer class.

$$\text{Accuracy} = \frac{\text{number of correctly predicted class}}{\text{number of total class}} \times 100\% \quad (9)$$

$$= \frac{19}{20} \times 100\%$$

$$= 95\%$$

V. CONCLUSION

Calculations are connected with presumptions, for example, topsy-turvy property of CXR and knobs speaks to are just lung disease knobs. This exploration has effectively built up an arrangement of tiny lung biopsy picture examination to recognize lung disease. The computerized picture preparing includes surface highlights extraction utilizing the Gray Level Co-Occurrence Matrix (GLCM) and Gray Level Run Length Matrix (GLRLM) technique and picture arrangement utilizing the Convolutional Neural Network (CNN) calculation.

Surface highlights are removed dependent on parameters of complexity, connection, vitality, and homogeneity, while tiny lung biopsy pictures are ordered into either disease or non-malignant growth class utilizing the counterfeit neural system calculation. The recently created framework is fit for grouping pictures with 93% precision on the preparation organize, and 97% exactness on the testing stage. These two outcomes demonstrate that this framework is appropriate to be executed for lung malignant growth location purposes.

VI. FUTURE SCOPE

In future, Developer extend this project with pattern classification using classifiers such as SVM classifier. As for further development segmentation process can be improved along with the lung nodule extraction methods where artificial intelligent methods can be used which ultimately increase the accuracy of the tested results.
ANN also needed to be continuing on lung nodule detection from blob area values which needed to be incorporated with further testing. According to the classification techniques, our work could be improved evaluating other classification algorithm as support vector machines, as well as improving the feature selection algorithm. It could be also very interesting to train the ANN in presence of noise.

REFERENCES

[1] American Cancer Society, “Cancer Statistics, 2005”, CA: A Cancer Journal for Clinicians, 55: 10-30, 2005, “http://caonline.amcancersoc.org/cgi/content/full/55/1/10”.

[2] B.V. Ginneken, B. M. Romeny and M. A. Viergever, “Computer-aided diagnosis in chest radiography: a survey”, IEEE, transactions on medical imaging, vol. 20, no. 12, (2001).

[3] D. Lin and C. Yan, “Lung nodules identification rules extraction with neural fuzzy network”, IEEE, Neural Information Processing, vol. 4,(2002).

[4] B. Zhao, G. Gamsu, M. S. Ginsberg, L. Jiang and L. H. Schwartz, “Automatic detection of small lung nodules on CT utilizing a local density maximum algorithm”, journal of applied clinical medical physics, vol. 4, (2003).

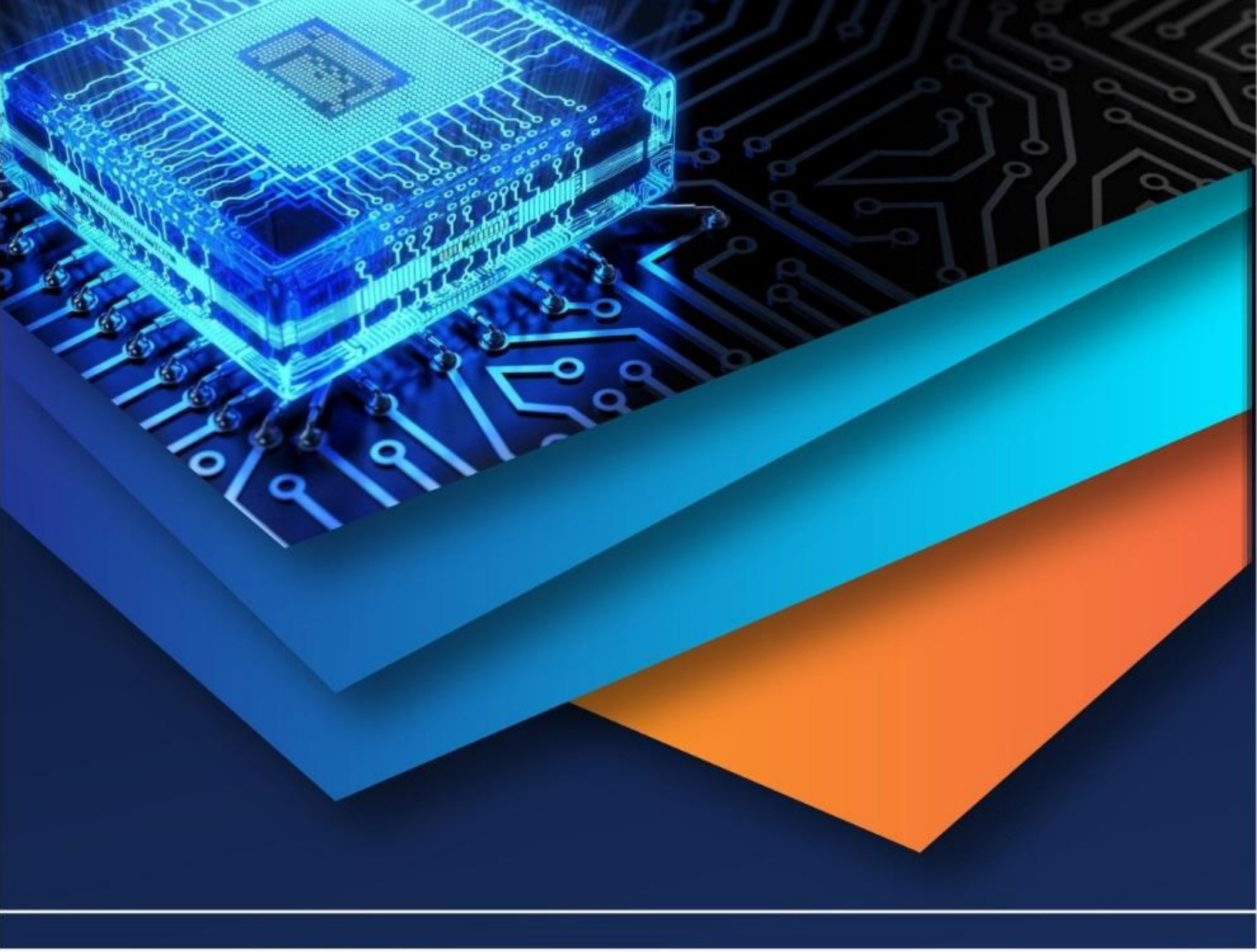
[5] A. El-Baz, A. A. Farag, PH.D., R. Falk, M.D. and R. L. Rocco,M.D.,“detection, visualization, and identification of lung abnormalities in chest spiral CT scans: phase I”, Information Conference on Biomedical Engineering, Egypt (2002).

[6] Linda G. Shapiro and G.C. Stockman., Computer Vision: Theory and Applications. 2001: Prentice Hall.

[7] The DICOM Standards Committee. DICOM homepage. http: //medical.nema.org/, September 2004.

[8] B. Magesh, P. Vijaylakshmi, M. Abhiram, “Computer aided Diagnosis System for identification and classification of Lessions in Lungs”, International Journal of Computer Trends and Technology- May to June Issue 2011.

[9] Rachid Sammouda, Jamal Abu Hassan, Mohamed Sammouda, Abdulridha Al-Zuhairy, Hatem abou ElAbbas, “Computer Aided Diagnosis System for Early Detection of Lung Cancer Using Chest Computer Tomography Images”, GVIP 05 Conference, 19-21 December 2005, CICC, Cairo, Egypt.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY