DETECTION OF LUNG CANCER FROM CHEST CT USING MACHINE LEARNING ALGORITHMS – SVM &KNN

P.HEMANTH KUMAR[1], R J S SUMANTH[2], U.THARAK NATH[3], SK.MOHAMMAD AYAZ[4]

[1] *Student, ECE, Vasireddy Venkatadri Institute of Technology, Andhra Pradesh, India*
[2] *Student, ECE, Vasireddy Venkatadri Institute of Technology, Andhra Pradesh, India*
[3] *Student, ECE, Vasireddy Venkatadri Institute of Technology, Andhra Pradesh, India*
[4] *Student, ECE, Vasireddy Venkatadri Institute of Technology, Andhra Pradesh, India*

ABSTRACT

Lung Cancer in today's world is one among the major outspread diseases which is the subject of maximum deaths every year. Early detection of lung cancer could boost the endurance rate. Biomedical image processing plays a vital role in the recognition of lung tumor using Computer Tomography (CT) scan images. In the detection Process lung CT images are subjected to image preprocessing followed by feature extraction and it is given to classification technique. Even though various image processing techniques and machine learning methods are applied for detection of lung cancer detection they are not accurate. There need to develop the best technique for the detection which improves the accuracy. In this paper we proposing two techniques SVM and KNN for the texture feature classification. These techniques comprises feature extraction, fusion using the patch base LBP(Local Binary Pattern) and DCT(Discrete Cosine Transform). The proposed techniques achieves better accuracy of 93% and 91% for Support Vector Machine and K-Nearest Neighbors respectively.

Keyword: - *Cancer; Diagnosis; LBP; CT; Machine Learning Techniques.*

1. INTRODUCTION

Lung Cancer is the second deadliest disease globally due to worldwide lack od diagnosis and proper treatment in the early stage. According to WHO report issued in 2020, the mortality rate was around 10 million due to cancer, while lung cancer mortality rate was around 2.1 million (20%)[1]. It is difficult to detect in early stage because it arises and shows symptoms in final stage in most number of cases. However the early detection and diagnosis can improve the survival rate. The abnormal cell growth results in a tumor in the lung, also know as carcinoma. Cancer spread out to adjacent organs through an infected cell in blood or lung tissue lymph fluid[3]. The most common indications of lung cancer are cough, respiratory infections, weight loss and chest pain[4]. Lung cancer is categorized into Small cell lung cancer(SCLC) and Non small cell lung cancer(NSCLC).The SCLC is most malignant cancer that occurs in 15%, while NSCLC occurs in remaining 85% which is further divided into as adenocarcinoma, large cell carcinoma and squamous cell carcinoma[1]. Out of this adenocarcinoma occurrence causing 38.5% of cases[5]. In lung cancer there are 4 stages where in stage 1 and 2 cancers limited to the lungs mostly and in stage 3 and 4 cancer infect to other adjacent organs.[6][7]. Different diagnosis techniques have been used during lung cancer detection such as MRI(magnetic resonance imaging), PET (positron emission tomography), CT (computed tomography), chest X-rays and more but CT scan images are used in most cases due to its robustness and low noise in tumor size finding[8][9].

2. LITERATURE REVIEW

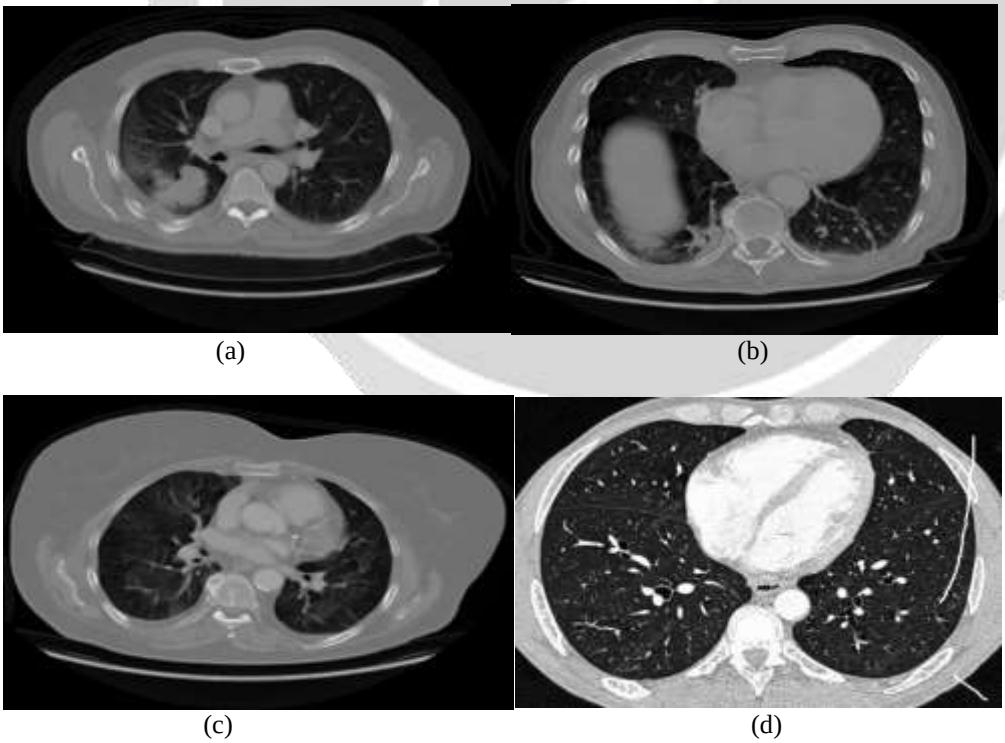
LBP is commonly used as a texture and appearance feature descriptor. LBP works only on gray scale images, a 3*3 patch is selected for LBP computation around the each pixel. The central value of the patch is updated according to the 8 neighbor values by comparing either in clock wise or anticlockwise direction, which is after converted into binary. Based on the central pixel value it is set to either 0 or 1. If the neighborhood value is less than the central value it is set to 0 otherwise 1[10].

New model was implemented based on image processing techniques in[11] to detect the lung cancer. Local Optimal Oriented Pattern (LOOP) for feature extraction followed by SVM classification of chest CT scan images. The performance of the model is analyzed by comparing with Local Binary Count, Local direction pattern and Local Binary Pattern. The LOOP model achieves an accuracy of 98.7% ANN based model; was developed for lung cancer detection by considering the gray scale image chest CT from the data set. Log-Gabor technique used for image enhancement, histogram equalizer has been evaluated and Marker- Controlled watershed segmentation for segmentation followed by Grey Level Concurrence Method used for feature extraction[12].

CT scan images were used for early lung cancer detection through a newly developed contextual clustering technique. The system segments nearly 1000 patients CT scans of LIDC IDRI -data set and extract features through GLCM. Then classified in terms of Accuracy, Sensitivity and Specificity [15]. The second and third LBP comparison analysis shows that the 3d LBP is better in accuracy 78% and by PNN it is 43% by evaluation of Japan Society of Computer Aided Diagnosis of medical images data set CT scan images.[17]. Figure 1 shows different types of lung cancer chest CT scan images.

3. RESEARCH METHODOLOGY

The technique has been developed to detect and diagnose early lung cancer, based on the feature extraction through LBP. The new technique applied to lung cancer CT scan image for global texture feature extraction which involves basically 3 main stages. First stage describes three patch LBP which is used for the texture features selection of CT chest scan Images and the second features are extracted through DCT using the prior stage. And it is combined to next stage. Finally, In the last stage the classification using machine learning classifiers SVM based on the polynomial kernel and KNN to classify the CT images of chest For classifying it into cancerous and von-cancerous. The proposed technique flow is presented in fig.2.



(a) (b)

(c) (d)

Fig.1. Chest CT scan images of Lung Cancer subtypes.(a) Adeno carcinoma (b) Large cell carcinoma (c) Squamous cell carcinoma (d) Normal

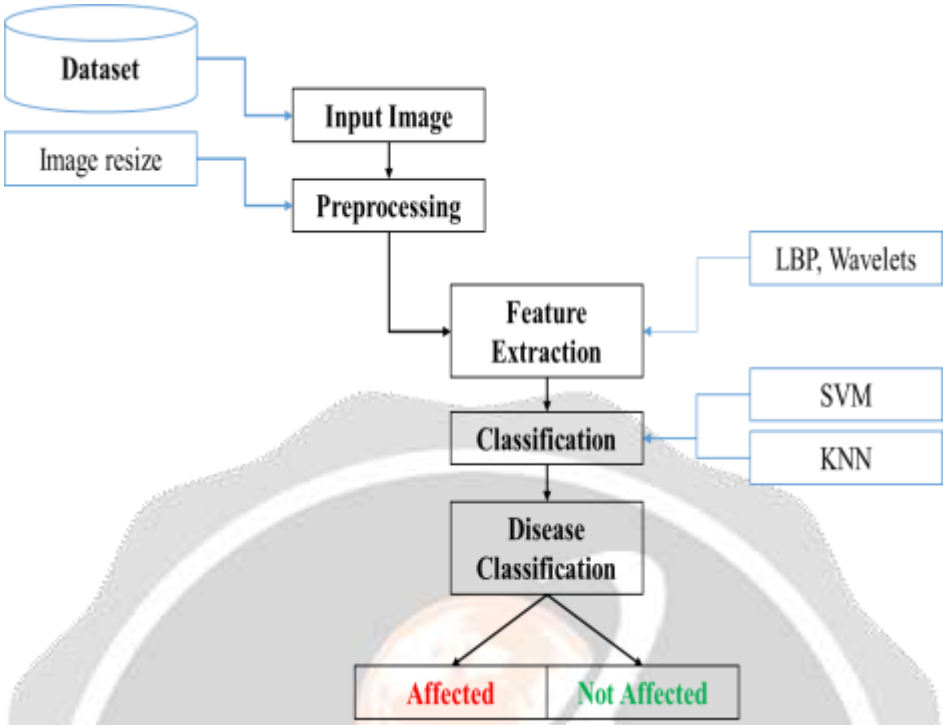
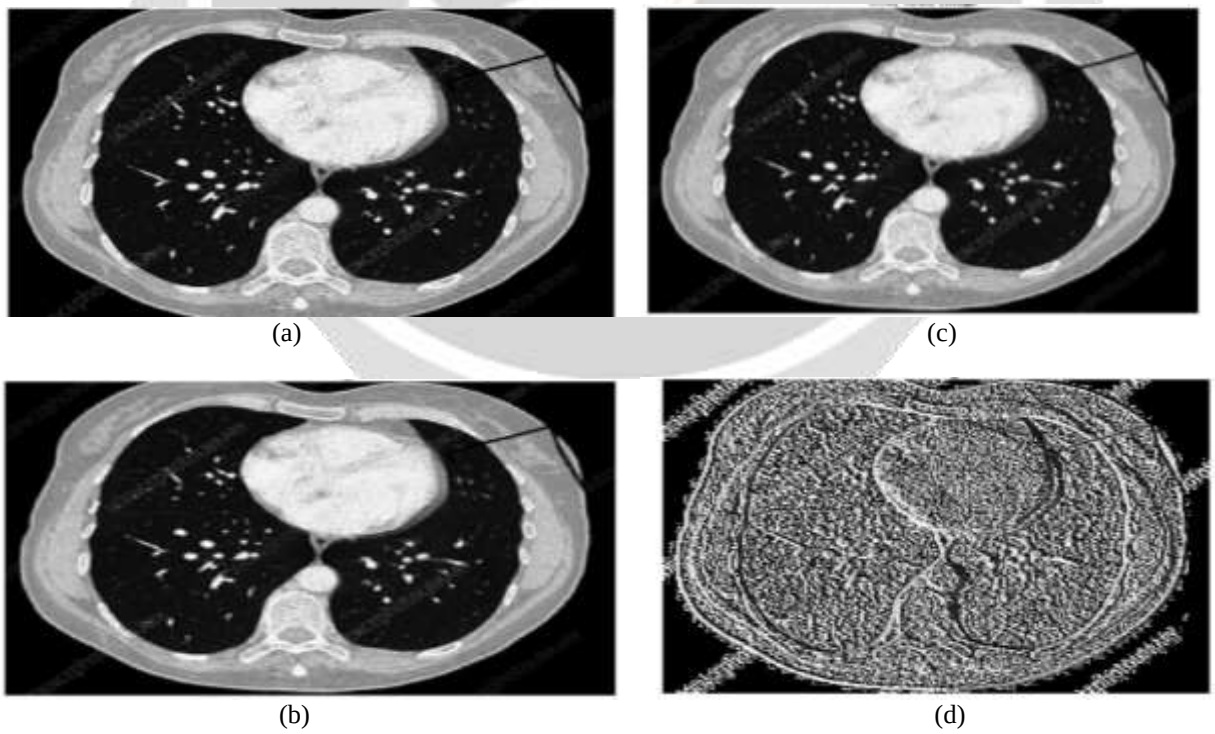


Fig.2 proposed research framework



(a) (c)

(b) (d)

fig 3. (a) Input Image (b) Resize Image (c) Grey scale Image (d) LBP image

4.SOFTWARE

In this project we used MATLAB 2015b version with Windows as an Operating system.

5. EXPERIMENTAL RESULTS

The proposed technique is evaluated on publicly available Chest CT scan images dataset. The techniques performance has been analyzed by 3 parameters called sensitivity specificity and Accuracy with state of art. The chest CT scan images contains three types of cancerous images as adenocarcinoma(ADC), Large cell carcinoma(LCC), Squamous cell carcinoma(SCC) and normal(NOR). Chest CT images have been evaluated for experimental analysis for each type of cancer and non cancer cells. 60% of CT images were used for training and remaining for the testing and validation. The results were calculated by taking input images of different cancer and non cancer by encoding texture based features using TLBP. The experiments results demonstrated that the proposed technique performance is better than the state-of-the-art techniques by achieving an average accuracy of 93% and 91% in the case of SVM and KNN. The recognition rate of lung cancer using SVM and KNN is shown in Table Ⅰ and Table Ⅱ. The comparision of the proposed technique represents most effective than existing techniques in lung cancer detection and diagnosis. The comprehensive analysis reported that SVM classifier detection rate in ADC, LCC, and SCC respectively are 91.66%, 91.33% and 89%. KNN classifier is implemented to overcome the SVM classifier limitatiom in detecting the SCC. KNN technique illustrates the detection rate in ADC, LCC and SCC ,respectively are 87%, 87% and 91.33%.

	ADC	**LCC**	**NOR**	**SCC**
ADC	59	2	0	4
LCC	3	61	0	5
NOR	0	0	68	0
SCC	12	14	0	47

Table 1.recognition rate of lung cancer using SVM

	ADC	**LCC**	**NOR**	**SCC**
ADC	62	4	0	9
LCC	3	65	0	5
NOR	0	0	71	0
SCC	9	5	0	51

4. CONCLUSIONS

This proposed system provides the real time application for the protection and detection of vehicles with the help of IOT notification wirelessly with raspberry pi pie camera and internet as its major components. It uses hybrid mechanisms such as Haar cascade algorithm for the identification and recognition of images that is captured by the pi-camera. The system is low cost, secure and highly efficient and thus can acts as a very good vehicle guard for the Sautomobile sector

5. REFERENCES

[1]. Amritha Nag,Nikhilendrao J N, Mrutyunjay Kalmath, "IOT Based Door Access Control Using Face Recognition",2018 3rd International Conference for Convergence in Technology (I2CT), pp 1-3

[2]. Prof.K.T. Jadhao and Prashanth Balraj Balla ", IOT Based Facial Recognition Security system",2018 Alamuri Ratnamala Institute of Engineering and Technology (ARIET) pp 1-4.

[3]. Prabal Deep Das, Sharmila Sengupta, "Proposing the systems to provide protection of vehicles against theft and accident",2016 IEEE Conference On Recent Trends In Electronics Information Communication Technology", pp 1681-1685.

[4]. S.Padmapriya & Esther Annlin KalaJames, "Real Time Smart Security System Using Face Detection and Recognition",2012 International Conference on Computer Communication and Informatics (ICCI-2012) , pp 1-6.

[5]. Mahesh R. P. Imdad R. , "IoT Based Embedded System for Vehicle Security And Driver Surveillance", Proceedings of the 2nd International Conference on Inventive Communication and Computational Technologies (ICICCT 2018), IEEE Explore Compliant - Part Number: CFP18BAC-ART; ISBN:978-1-5386-1974-2.

[6]. Vivek K. S., Soumitra M., and Harshit M.," Car Security using Internet of Things", 1st IEEE International Conference on Power Electronics Intelligent Control and Energy Systems (ICPEICES 2016), 978-1-4673-8587-9/16/31.00 ©2016 IEEE.