International Journal of Advanced Trends in Computer Applications
www.ijatca.com

Scale-invariant feature transform based Lesion Detection on CT Lung Images

[1]Gurpreet Kaur, [2]Navneet Kaur
[1] *M.tech (Scholar), CSE*
Punjabi University, Patiala, India
[3] *Assistant Professor, CSE*
Punjabi University, Patiala, India

[1] *kaurgurpreet692@gmail.com*, [2] *mavi_navneet@yahoo.com*

Abstract: *As the image processing field grows day by day, researcher moves towards bio medical field to emerge new techniques and to diagnose various medical diseases using automated image processing algorithms. One of them is Lung lesion detection also known as Cancer Detection. Many researchers has worked on lung lesion detection. But successful interpretation mainly depends on the feature extraction, so it is very important and crucial step in lung lesion detection. So, this paper has proposed a unique method based on the combination of SIFT feature extraction method with two classification algorithms i.e. neural network and Support Vector Machines Algorithm. The proposed algorithm is based on 16 CT images for implementation of the algorithm. From result evaluation it has been seen that NN outperforms than SVM in terms of classification value.*

Keywords: *Lesion Detection, SVM, NN, Segmentation, Classification, Feature extraction.*

I. INTRODUCTION

Biomedical image processing has encountered sensational development and has been an interdisciplinary examination field drawing in ability starting physics, applied mathematics, medicine, engineering, statistics, computer sciences, as well as biology [1]. Computer-assisted diagnostic processing has turned into a vital piece of clinical program/procedure. It is joined by a flow of innovative development of high innovation and utilization of different imaging modalities, more difficulties emerge; for example, in what way to process as well as examine a significant volume of pictures so that astounding data can be delivered for sickness findings and treatment[2,3].

Medical technology has made the doctors to see the interior body part of an active living being [4]. This one correspondingly supports doctors in numerous surgical procedures similar to MRI scan, X-ray, Ultrasound as well as CT scanner. By utilizing CT scanner, internal parts could be observed without giving any pain to the patient [5]. MRI gets the signal from the body magnetic particles and with the computer [6,7]. For analysing remote sensing the data image processing techniques are also developed that may be modified to analyse the outputs of medical imaging systems to get advantage to analyse symptoms of the patients with ease [8,9]. The lung lesion occurs when anomalous tissue discovered on or else inside an individual's lung. This specific tissue in maximum circumstances is impaired as well as may possibly be in the course of healing, then again there are probabilities in which it possibly will remain ceaseless. These specific groups gathering of tissues normally provide the impression as white, spherical shades on x-rays or CT scan [10,11].

In this proposed work, we have utilized two techniques hybridization which is of SIFT algorithm with NN and SVM for lung lesion detection. Here, SIFT algorithm is utilized for feature extraction purpose while NN and SVM are utilized for training, testing and classification and also compare the results of both simultaneously for enhancing and getting better result from earlier in lung lesion disease detection framework.

Remaining paper is organised as: Section 2 includes the literature survey, Section 3 contains the various techniques used in proposed work, Section 4 discusses the proposed methodology, Section 5 discusses the results and implementation section and finally section 6 contains the conclusion and result part.

II. RELATED WORK

Yang Song presented an automated method to detect and characterize the primary lung tumor and disease in regional lymph nodes in thoracic FDG PET-CT images from NSCLC studies. They propose a context driven approximation method to distinguish between lesions and soft tissues, and between lung tumors and abnormally mph nodes. New sparse representation and multi-atlas models were designed with additional constraints to improve the labeling performance. Patch-based lesion detection and object-based lesion characterization were designed based on approximation with region and image-level contexts. They evaluated a method on clinical dataset and showed that the method outperformed the state-of-the-art approaches.[12]

Shanti Mahesh recognized lesions and its qualifications from a selected MRI scan of a brain image and compute the world of each lesion by thresholding. Thresholding approach section scalar pictures by generating a binary partitioning of the image intensities. Otsu's technique is employed to mechanically perform intensity primarily based image thresholding. The quantitative chemical analysis of MRI brain lesions permits helpful key info concerning the lesions. Segmentation is finished on basis of threshold, as a result of that whole image is regenerate into binary image [13]**.**

Mahound conferred a comprehensive review of the ways and techniques accustomed find neoplasm through MRI image segmentation. The paper concludes with a terse discussion and provides a direction toward the coming trend of a lot of advanced analysis studies on brain image segmentation and growth detection [14].

U. Javed proposed a new way of working out best loads for the computed features. This recommended approach is usually tested with CE CT Lung pictures. Simulation effects and research exhibited that his or her recommended method indicates greater distinction exactness than the conventional SVM. It is multilayer design are designed for complicated difficulties. In line with the entropy estimation, theory-based finding out is conducted in your community with just about every neuron. Sensory variables and cable connections that correspond to minimum entropy tend to be adaptively arranged for every single neuron [15]**.**

Donia Ben Hassen proposed an approach for lesion detection from chest radiography. They have described the significance of precise segmentation as a preprocessing step in a CAD scheme. Then using the forward stepwise selection method,anappropriate combination among 118 features has been recognized. The main idea is to find a set of features that enables a CAD not to distinguish between normal lesions and abnormal ones but to specify its nature if this lesion has an infection [16].

U.Javed proposed a new way of working out best loads for the computed features. This recommended approach is usually tested with CE CT Lung pictures. Simulation effects and research exhibited that his or her recommended method indicates greater distinction exactness than the conventional SVM. It is multilayer design are designed for complicated difficulties. In line with the entropy estimation, info theory-based finding out is conducted in your community with just about every neuron. Sensory variables and cable connections that correspond to minimum entropy tend to be adaptively arranged for every single neuron [20].

V.V.Thakre, presented a whole new technique of calculations of plot size of a square micro strip plot antennas employing (ANN). In this papers a whole new tactic is usually recommended to development inset nourish micro strip antenna with slot machines inside it to raise the antenna bandwidth. This papers relates to the planning of slotted micros rip antenna on the substrate of fullness 1. 588 mm giving wide band qualities employing ANN. Your illustrated plot antenna gives increased bandwidth as compared with antenna devoid of slot machines with the same actual physical size. In our function a good Synthetic Sensory Multilevel (ANN) design is usually formulated for you to analyze the bandwidth with the case antenna. The method of Moments (MOM) structured IE3D computer software has been utilized to generate training and examination information for the Synthetic neural multilevel [21].

III. BASIC CONCEPTS OF PROPOSED ALGORITHMS

3.1 SIFT (Scale Invariant Feature Transform)Algorithm

Scale Invariant Feature Transform (SIFT) was proposed by David Lowe that has the capacity to distinguish and depict neighborhood picture elements positively[17].

A. Scale space local extreme detection

The Difference of Gaussian (DOG) is computed as of images are:

$$D\,(a1, b1, \sigma) = (G\,(a1, b1, k\,\sigma) - G\,(a1, b1, \sigma)) * J\,(a1, b1) = L\,(a1, b1, k\,\sigma) - L\,(a1, b1, \sigma$$

(1)

Where k is a constant multiplicative factor in scale space that is used for changing the scale and a1, b1 are the coordinates of a pixel in image J.

B. Key-point localization

To detect the importance points, DOG pictures are utilized also local maxima as well as local minima are computed across different scales.

C. Orientation assignment

A Gaussian smoothed image $L1$ is selected using the scale of a particular key-point. On behalf of a Gaussian smoothed picture $L1\ (a,b)$,mmagnitudes$(m(a,b))$ and orientation $(\theta(a,b))$ are calculated as

m (x,y)= $\sqrt{\left(L_1(x+1,y)-L_1(x-1,y)\right)^2+\left(L_1(x,y+1)-L_1(x,y-1)\right)^2}$ (2)

$$\theta\ (x,y) = \tan^{-1}\left(\frac{(L(x,y+1)-L(x,y-1))}{(L(x+1,y)-L(x-1,y))}\right) \quad (3)$$

D. Key-point descriptor

The picture gradient magnitudes and introductions, with respect to the significant introduction of the key point, are inspected inside a 16X16 locale around every key-point.

E. Trimming of false matches

Compute the predominant orientation QP and length lP of the matching, and keep the matching pairs whose orientation µ and length ` are within predefined tolerances ϵQ and ϵP so that $|Q\ - QP|\ <\ \epsilon Q\ and\ |l - lP|\ <\ \epsilon l.$

3.2 SVM (Support Vector Machines)

Support Vector Machine (SVM) also called Support Vector Networks are supervised learning models that are usedto analyzes data and recognizes patterns [18].

Algorithm1

Load the model
Classify the dataset
Train SVM datase
Change parameters values
Perform validation
Save model
Retrieve results

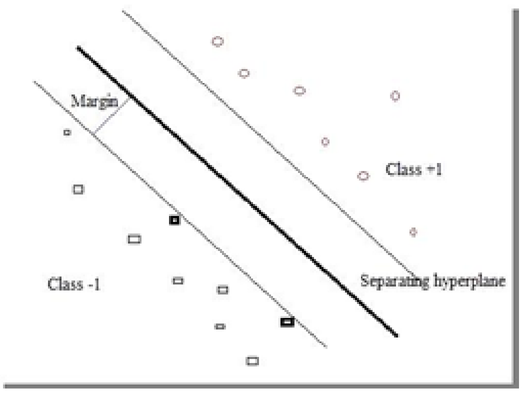

Fig.1. SVM Working

To minimize the error for unseen patterns, SVM finds the optimal separating hyper plan. Consider the problem of unravelling the set of training vectors belonging to two separate classes.

$$x_1, x_2, \dots . x_n \quad (4)$$

Which are vector in $RD.$

We consider a decision function of the following form:

$$yx = wT\emptyset x + b \quad (5)$$

Attached to each observation, $x$is a class label,$t_i \in -1, +1.$ Without loss of generality, we must construct a decision function such that,$yx_i > 0$ for all *i* such that $t_i = +1$ and y$x_i < 0$ for all I such that $t_i = -1.$ We can combine these requirements by stating,

$$t_i y x_i > 0 \forall i \quad (6)$$

The idea is to extend it to multi-class problem is to decompose an M-class problem into a series of two-class problems.

3.3 NN (Neural Network)

Machine learning algorithms facilitate a lot in decision making and neural network has performed well in categorization purpose in medical field. Most popular techniques among them are neural network [19].

A neural network is a connected node set that depends on human neuron system. As shown in figure below, number of inputs i.e. $x_j = 1, \dots M_j$ which is pass to the neural network by the weights. The node in the circles represents the function of summation of weighted incoming signals and it can be shown as:

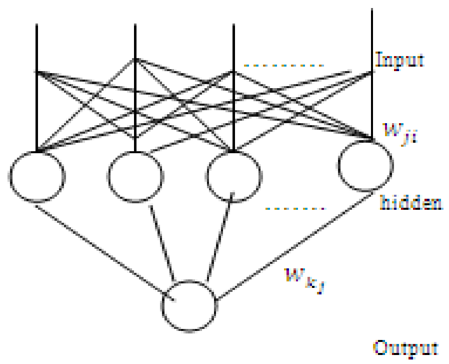

Fig.2. Neural Network

$$o_k = f(a) \quad (7)$$

$$a = \sum_{j=1}^{M_j} x_j w_{kj} \quad (8)$$

$w_{kj}$is the weight among the input layer j and the hidden layer k.o_k is the output from the node and f(a) is the activation function of the node. The activation function given in the logistic function is shown as:

$$f(a) = \frac{1}{1-e^{-a}} \qquad (9)$$

Neural network training has the function to search the weight set that minimizes the cost function on given data set i.e. $D = \{x_m, t_m\}$, $x_m$is the input vectors and t_m is the class (0 and 1) for the nth training case. The activation function is always between 0 and 1. The output of the neural network can be defined as:

$$y_n = y(x_n; w) \equiv P(t = 1 \vdots x_n, w) \qquad (10)$$

Therefore, the weighted function is:

$$P(D \vdots w) = \prod_n y n^{tn} (1 - y_n)^{1-t_n} \qquad (11)$$

The error function in neural network is the entropy function of the above equation:

$$E(D \vdots w) = -\sum_n t_n \log(y_n) + (1 - y_n) \qquad (12)$$

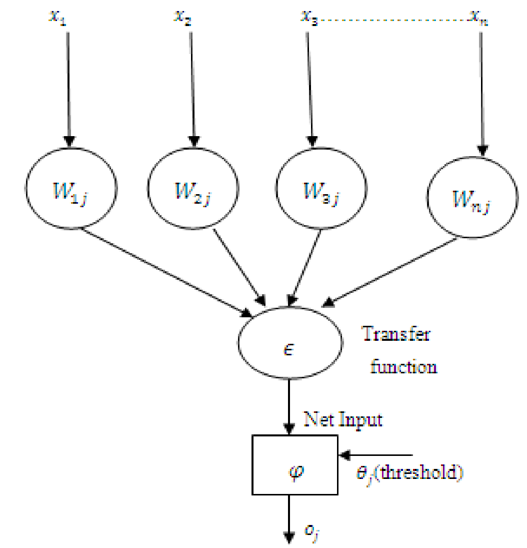

Fig.3. Artificial Neural Network

Algorithm 2

```
net.trainFcn  = 'trainlm'
[net,tr]  = train(net,...)
net.trainFcn = 'trainlm' sets the network trainFcn property.
[net, tr] =
train(net, ...) trains the network with trainlm.
```

IV. PROPOSED METHODOLOGY

The parts of the image containing the lesions normally have more intensity than the other portion and we assume the area and length of the tumour in the image and the code goes through the following steps.

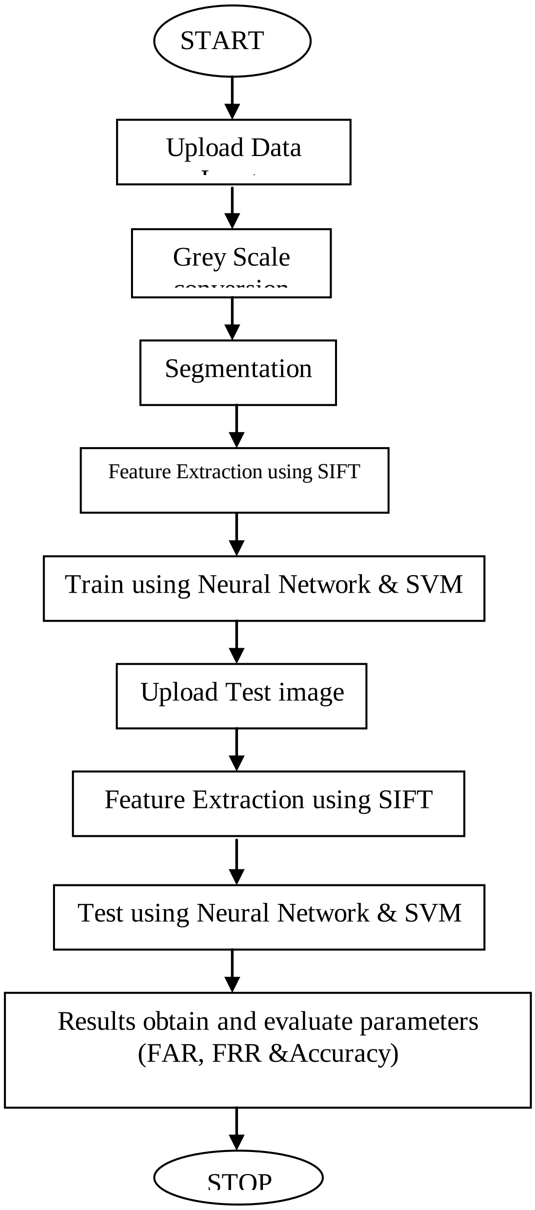

Fig.4. Proposed Flowchart

i. On initializing we upload the data image file.
ii. Then grey scale conversion of image is processed.
iii. Once image is in grey scale, it is segmented in different yet equal parts.
iv. Once it is segmented, SIFT is applied for feature extraction.
v. Then we train the data utilizing Neural Network and Support Vector Machine.
vi. After completion of training we will upload test image.
vii. Again SIFT is used for feature extraction from the chosen image.
viii. Then we test the extracted features separately once by neural network and then by SVM. We will obtain result separately.
ix. Then evaluate results on the basis of these specific parameters like, FAR, FRR, and Accuracy. And also compare which technique is better NN or SVM.

V.RESULTS AND IMPLEMENTATION

The proposed lesion detection system has different category panels first one is feature extraction panel in which there are various buttons to upload the dataset for < 10, upload the dataset lesion CAT>20<30, train data using neural network and SVM training. In second and last panel is test panel in which we have some buttons of different category given as: upload test image as soon as it is uploaded simultaneously SIFT algorithm is applied for feature extraction and feature optimization respectively, then testing is done by utilizing neural network and SVM algorithm and by testing we get three parameters as result which are: FAR, FRR and Accuracy.

Below plotted graph shows the comparison between neural network and Support vector machine using parameters FAR, FRR and Accuracy. In the comparison it shows that neural network gives better performance as compared to support vector machine. Its obtained value is also shown by beneath given table. It has been shown that average value of FAR, FRR and Accuracy has been shown in below graph. Here red color is for SVM and blue color is for NN.

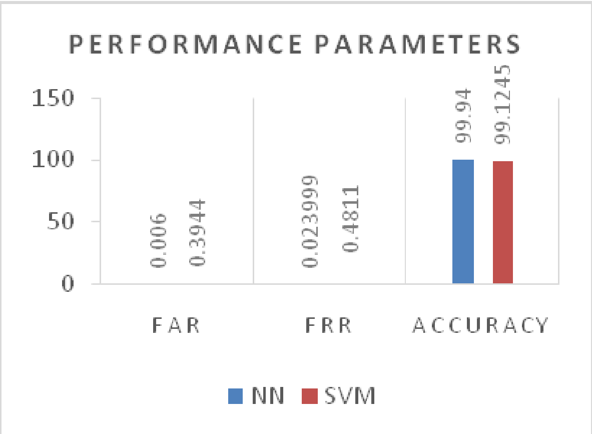

Fig.5. Comparison graph between various parameter of NN and SVM

Table 1. COMPARISON BETWEEN NEURAL NETWORK AND SUPPORT VECTOR MACHINES

Technique	FAR	FRR	Accuracy
Neural Network	0.006	0.023999	99.94
Support Vector Machine	0.3944	0.4811	99.1245

VI. CONCLUSION AND FUTURE SCOPE

In our proposed work, we have proposed a system in which we utilized hybridization of SIFT algorithm with NN and SVM. Here, SIFT algorithm is utilized for feature extraction purpose while NN and SVM are utilized for training, testing and classification. In this, we have compared the results of neural network with support vector machine on the basis of parameter FAR, FRR and Accuracy. The FAR value of NN is 0.006 and SVM is 0.3944, FRR value of NN is 0.023999 and SVM is 0.4811 and accuracy value in NN is 99.94% and SVM is 99.12%. In proposed work combination of SIFT +NN and SIFT +SVM has been utilised to show whether SVM+SIFT works well or SIFT+NN works well and in the end it has been concluded that SIFT+NN worked well in terms of FAR, FRR and accuracy. The future scope of this work is that we can use hybridisation of other techniques similar to SIFT i.e. ICA for feature extraction with Neural Network or K-mean clustering and also compare their results for better output.

REFERENCES

[1] Y. Song, W. Cai, S. Eberl, M. J. Fulham, and D. Feng, “Automatic detection of lung tumor and abnormal regional lymph nodes in PET-CT images,” J. Nucl. Med., vol. 52, no. Suppl. 1, p. 211, 2011.

[2] H. Cui, X. Wang, and D. Feng, “Automated localization and segmentation of lung tumor from PET-CT thorax volumes based on image feature analysis,” in Proc. IEEE EMBC, 2012, pp. 5384–5387.

[3] H. Ying, F. Zhou, A. F. Shields, O. Muzik, D. Wu, and E. I. Heath, “A novel computerized approach to enhancing lung tumor detection in whole-body PET images,” in Proc. IEEE EMBC, 2004, pp. 1589–1592.

[4] C. Lartizien, S. Marache-Francisco, and R. Prost, “Automatic detection of lung and liver lesions in 3-D positron emission tomography images: A pilot study,” IEEE Trans. Nucl. Sci., vol. 59, no. 1, pp. 102–112, Feb. 2012.

[5] Y. Song, W. Cai, H. Huang, X. Wang, S. Eberl, M. Fulham, and D.mFeng, “Similarity guided feature labeling for lesion detection,” in MICCAI. Heidelberg, Germany: Springer, 2013, vol. 8149, LNCS, pp. 284–291.

[6] D.Hellwig, T. P.Graeter, D.Ukena, A. Groeschel, G. W. Sybrecht,H. J. Schaefers, and C.M. Kirsch, “18F-FDG PET for mediastinal staging of lung cancer:Which SUV thresholdmakes sense,” J. Nucl.Med., vol. 48, pp. 1761–1766, 2007.

[7] C. Lartizien, M. Rogez, A. Susset, F. Giammarile, E. Niaf, and F. Ricard, “Computer aided staging of lymphoma patients with FDG PET/CT imaging based on textural information,” in Proc. Int. Symp. Biomed. Imag., 2012, pp. 118–121.

[8] I.Jafar, H. Ying, A. Shields, and O. Muzik, "Computerized detection of lung tumors in PET/CT images," in Proc. IEEE EMBC, 2006, pp. 2320–2323.

[9] Y. Cui, B. Zhao, T. Akhurst, J. Yan, and L. Schwartz, "CT-guided automated detection of lung tumors on PET images," in Proc. SPIE Med. Imag., 2008, vol. 6915, p. 69152N.

[10] C. Ballangan, X.Wang, S. Eberl, M. Fulham, and D. Feng, "Automated detection and delineation of lung tumors in PET-CT volumes using a lung atlas and iterative mean-SUV threshold," in Proc. SPIE Med. Imag., 2009, vol. 7259, p. 72593F.

[11] J. Gubbi, A. Kanakatte, T. Kron, D. Binns, B. Srinivasan, N. Mani, and M. Palaniswami, "Automatic tumour volume delineation in respiratory- gated PET images," J. Med. Imag. Radiat. Oncol., vol. 55, pp.65–76, 2011.

[12] Yang Song, WeidongCai, Heng Huang, Xiaogang Wang, Yun Zhou, Michael J. Fulham, and David Dagan Feng, "Lesion Detection and Characterization With Context Driven Approximation in Thoracic FDG PET-CT Images of NSCLC Studies" ieee transactions on medical imaging, vol. 33, no. 2, february 2014.

[13] Shanthi Mahesh, "A New Approach For An Improved Multiple Brain Lesion Segmentation", International Journal Of Innovative Research In Computer And Communication Engineering (An Iso 3297: 2007 Certified Organization), Vol2, Issue.4, 2014.

[14] Mohammad, "Tumor Brain Detection Through Mr Images: A Review Of Literature", Journal Of Theoretical And Applied Information Technology, Pp.387-404, 2014.

[15] U. Javed, M. M. Riaz, T. A. Cheema and H. M. F. Zafar., "Detection of Lung Tumor in CE CT Images byusing Weighted Support Vector Machines", IEEE, pp. 113-116, 2013.

[16] Donia Ben Hassen and HassenTaleb, "Lesion Detection in Lung Regions that are segmented Using Spatial Relations" 2012 International Conference on Information Technology and e-Services ©2012 IEEE.

[17] http://docs.opencv.org/master/da/df5/tutorial_py_sift_intro.html.

[18] www.support-vector-machines.org/

[19] Z. Daniele, H. Andrew, J. Nickerson, "Nuclear Structure in Cancer Cells," Nature Reviews Cancer, Medical School, vol. 4, no. 9, pp. 677-87, USA, Sep. 2004. (Pubitemid 39215066).

[20] U. Javed, M. M. Riaz, T. A. Cheema and H. M. F. Zafar., "*Detection of Lung Tumor in CE CT Images byusing Weighted Support Vector Machines*", IEEE, pp. 113-116, 2013.

[21] V. V. Thakare, P. Singhal, "*Neural network based CAD model for the design of rectangular patch antennas*," JETR, Vol. 2(7), 2010.