BLB (Brain/Lung Cancer Detection and Segmentation and Breast Dense Calculation)

Shobanadevi K[1] Anupriyaa.M[2] Pavithra.R[3]

Prof. Aarthi.C[4]

[1,2,3]*Department of ECE, Sengunthar Engineering College, Tiruchengode,(India)*

[4]*Assistant Professer, Department of ECE, Sengunthar Engineering College, Tiruchengode,(India)*

ABSTRACT

Cancer Diagnosis and prediction have been one of the most important challenges faced by mankind in the last few decades. Accurate detection of cancer would facilitate saving millions of lives. Therefore, our research presents Brain/Lung/Breast (BLB) automated detection system. We precisely predict the occurrence of cancer and segment the expected region of tumor/cancer in MRI/CT scan/Mammography. This system proposes different classification techniques including Support Vector Machine (SVM), Normal MRI, Extra Trees and Convolutional Neural Network (CNN). SVM achieved 98% accuracy in classifying brain tumor and Normal in MRI images. Extra Trees achieved 95.38% accuracy in classification of Breast Dense in mammography images. Convolutional Neural Network (CNN) deep learning accomplished acceptable results in lung cancer detection. The presented system consists of (Preprocessing/Post processing, Segmentation, Feature extraction, and Classification). This approach was applied on different datasets including (MIAS) for breast, (LUNA) for lung and (BRATS2015 and Brats2017) datasets for Brain Index Terms—SVM, CNN, Brain Tumor, Lung Cancer, Breast Dense.

I. INTRODUCTION

Cancer is one of the most harmful and critical diseases which leads to high rate of death. Radiologists detect cancerous regions visually from different types of scans. The accuracy of this process depends on the experience of radiologists. Many conflicts may occur in diagnoses and false prediction of cancer. A Computer-Aided Detection (CAD) system is required as an alternative solution. The proposed system can be considered as a reference point for radiologists that serve this point of research. Biomedical imaging is concerning with medical digital images. It use image processing to solve medical problems [1][2]. Meanwhile image processing concern with digital images to extract useful information from it. It is involved in different other topics such as Land-use planning [3], Character recognition [4][5] [6], coin classification [7].

A brain tumor is a mass or growth of abnormal cells in the brain [8]. According to [9] there are two types of tumor which are LGG - Low Grade Glioma (Grade I which is benign - looks like normal brain cells and grows slowly, Grade II which is Malignant- looks less likely to the normal brain cells) and HGG - High Grade Glioma (Grade III and Grade IV both of these grades are malignant and grows quickly). According to

[10] the average age of brain tumor identification is 59 years old however they most typically occur in children and older adults. Cancer segmentation is important in calculating tumor size and determining its location. Magnetic resonance imaging (MRI) [11] is a widely used technique for visual diagnosis. MRI provides significant information about the anatomy of the brain. MRI is used in our proposed system for Brain cancer classification.

A lung nodule is considered to be the first step in identifying lung cancer. There are two main types of nodules: malignant (cancerous) and benign (noncancerous). Smooth, round nodules are more likely to be benign, whereas irregular shapes are more likely to be cancerous and they are located in the upper portions of the lung. The nodule is defined as a spot on the lung which has a diameter. On the off chance that an anomaly is seen on a Ct scan is bigger than three centimeters, it is more likely to be cancerous[12]. Over 90% of nodules that are smaller than two centimeters (around 3/4 inch) in diameter are benign.[13] So, nodules can be an early sign for prediction of cancer.

Breast Dense is a common term which indicates the amount of dense tissue compared to the amount of fatty tissue in the breast on a mammogram [14]. According to MIAS dataset [15], there are three classes of dense: f - lowest grade.g-intermediate grade, d - highest grade. Some classes of dense could hide cancerous cells underneath them which may trick some doctors and lead to late breast cancer diagnoses. Calculation of breast dense level can vary from doctor to another. In order to eliminate this confusion, the system proposes a classification method for measuring breast dense through Mammography imaging. This paper is structured as the following: Section 2 will address the related works, section 3 for the proposed methodology, section 4 shows results and experiments, section 5 will address the conclusion, se This section will include three subsections which represent the related work for Brain, Lung and Breast separately each of which describes the workflow of the mentioned papers below.

II.RELATED WORKS

This section will include three subsections which represent the related work for Brain, Lung and Breast separately each of which describes the workflow of the mentioned papers below.

1.1. Related works for Brain tumor

Padma et al Introduced an algorithm for brain tumor detection. MRI image is taken as an input for the system which consequently enters the pre-processing stage to remove noise using a median filter and then applying threshold followed by connected components labeling, Sobel filter is used for edge detection thus the tumor is segmented.

Telrandhe et al Proposed an algorithm for tumor detection and segmentation. The MRI image enters the pre-processing stage such as brightness, threshold, and Filtering, then the image is passed to the skull masking phase which finds out the ROI (Region of interest). Morphological oper-ations are done in addition to drawing a line in the center of the stripped skull dividing the skull into two equal parts. One part is a test image and other is a reference image. As well as the K means segmentation followed by edge detection for tumor segmentation. Finally, the classification stage using SVM classifier.

Hwan-ho Cho et al proposed an automated system to predict (LGG) Low-grade glioma and (HGG) High-grade glioma in the brain. 45 radio mics features based on histogram, shape and gray-level co-occurrence matrix (GLCM) were extracted from FLAIR, T1, T1-Contrast, T2 images to measure the grade of Gliomas. Classification process was achieved through a 10-fold cross-validation among 54 LGG and 54 HGG images obtained from the BRATS 2015 dataset . And finally, the proposed approach was able to successfully achieve an accuracy of 89.81%.

N. Varuna Shree et al introduced a method of identification and classification of the brain tumor in MRI images. The author focused on noise removal techniques, extraction of gray-level co-occurrence matrix (GLCM) features and also discrete wavelet transform (DWT) for extracting wavelet coefficients. The classification was done through a Probabilistic neural network (PNN). This approach has achieved nearly 100% accuracy in identifying normal and abnormal tissues.

1.2. Related works for Lung cancer

Zhou et al , CT images of 70 cases of patients with pulmonary nodules are collected as experimental samples, which are marked by radiologists, with a size of 512 and a thickness of 2 mm. They are composed of 2232 CT images from 38 cases of patients with solitary pulmonary nodules (SPN), 17 cases of patients with vascular adhesion pulmonary nodules (VAPN), and 15 cases of patients with pleural adhesion pulmonary nodules (PAPN), respectively. After extracting the region of interest 42 features used including shape features, intensity and Texture feature passing by reduction to fit in SVM Model.

Gupta et al, presented a CAD scheme for automatic detection of multi-size pulmonary nodules in thoracic CT scans. The achieved training sensitivity is 85.6% with only 8 FPs/scan for the whole LIDC/IDRI dataset. The CAD scheme achieved an average sensitivity of 68.4% with only 8 FPs/scan on multiple independent heterogeneous test sets. Three Formulated subalgorithms modules done to detect different sizes of nodule candidates ($\geq$ 3 mm and < 6 mm; $\geq$ 6 mm and < 10 mm; and $\geq$ 10 mm). Each subalgorithm module included a multistage flow of rule-based thresholding and morphological processes

Elmar Rendon-Gonzalez and Volodymyr Ponomaryov applied several masks for thresholding. Suspicious Regions of Interest (ROI) are calculated using Hounsfield Units (HU). During feature extraction, numerous features are calculated in order to restrict the suspicious zones. Shape Features are extracted: area, Eccentricity, circularity and fractal dimension.Texture features used such as mean,variance, energy and entropy. Finally, Support Vector Machine (SVM) algorithm is employed in classification stage. The accuracy achieved with 79% for both train and test.

1.3. Related works for Breast dense and cancer

Vipul Sharma and Sukhwinder Singh introduced a system to classify breast dense which mainly depends on Region of interest (ROI) extraction which contains tissue pattern only excluding the pectoral muscle. Statistical features are extracted for the classification of breast dense. Classification was evaluated using Naive Bayes and KNN on MIAS dataset .

Samual H. Lewis and Aijuan Dong introduced a system for extracting the region of interest through marker-controlled watershed segmentation and morphological analysis applied on Mammographic Image Analysis Society (MIAS) dataset. Using preprocessing, segmentation generation, the location of foreground markers, computation of foreground markers and watershed segmentation, the obtained accuracy was 90%.

III.METHODOLOGY

The proposed methodology represents an automated system for Brain/Lung cancer detection and segmentation and Breast Dense calculation. The system overview is shown in Fig.1 represents the process flow for the Brain MRI image which passes by the feature extraction, the classification which identifies the image either tumor or not accordingly chooses the appropriate path, in case of tumor the image enters the post-processing phase to be prepared for the segmentation process. Considering Lung flow, the Lung CT scan image proceeds to the pre-processing step, segmentation where it develops an ROI, feature extraction phase to process down the extracted ROI, finally to the classification state. Considering Breast flow the mammography image enters segmentation, feature extraction, and classification

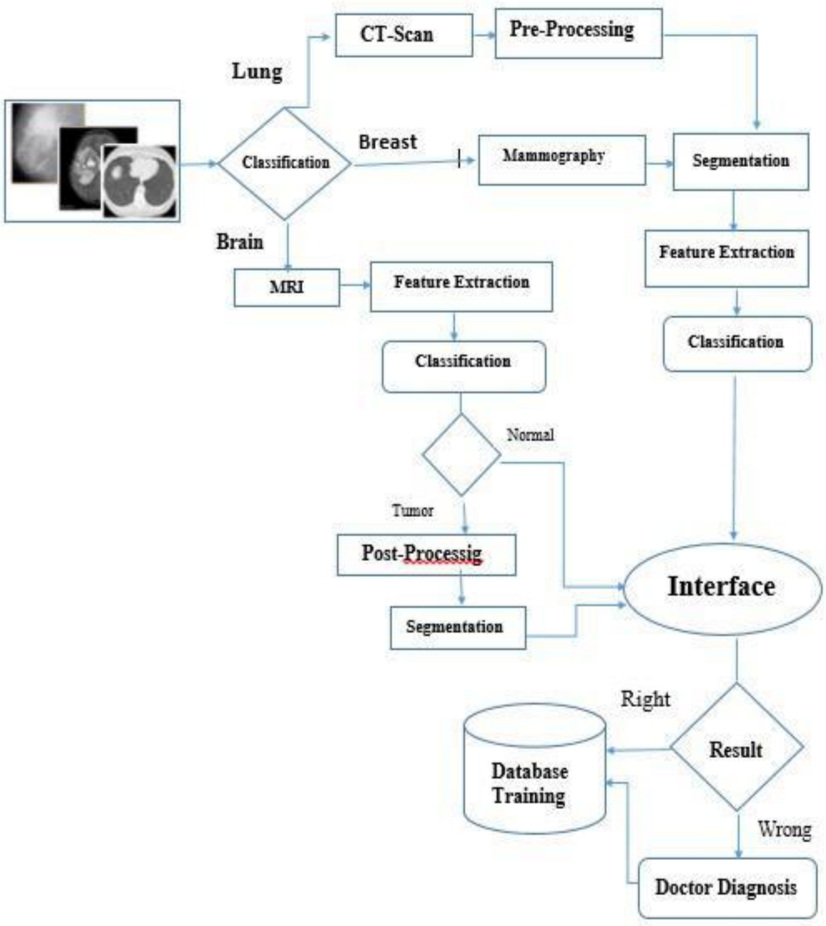

Fig. 1 System overview and flow

ISBN : 978-93-87793-75-0

2.1. Preprocessing

Preprocessing phase is required to remove the artifacts/noise from the image by applying some filters according to the noise found in the image. Gray pixel values are converted to House Field unit using Equation 1:

$$HU = (\text{Gray Value} * \text{slope}) + \text{intercept} \qquad (1)$$

Slope and intercept are given from raw data of images. After getting new values of house Field unit the lung can be extracted and the noise is excluded according to HU Value where the HU value of lung is -600. According to mhd formats pixel values should be converted from world co-ordinate to voxel co-ordinate to get the slice number, x and y position of the nodule. Lung Pre-Processing was tested on both Luna dataset and Kaggle competition data science bowl dataset and good results were achieved in viewing lung window with original slices of each patient.

2.2.MRI And Mammography Image: (Brain and Breast):

Another Features used for both Brain and Breast are nodules attached to the blood vessels. Closure operation with a disk of radius 10mm was the best fit to keep nodules attached to the lung wall. After that, the small holes inside the binary image are filled. Consequently, the image is multiplied by the mask to get segmented lungs. Finally, a tZreshold of -420 HU is used to minimize the values of lung walls to zero to get the region of interest .

2.2.1.Segmentation of Brain Tumor in MRI Images:

For MRI Images after getting features extracted, a Morphological open operation is applied with constant parameters which were the best fit to remove the artifacts from images. A median filter is applied for smoothing and removing the unwanted noise. After that, an average filter is applied for eliminating the pixels which do not contribute to the representation of its surroundings. A Gaussian filter is used to remove the fine details which are originally found in the image. After that unsharp masking filter is used to enhance the edges. Structuring Elements (Erode and Dilation) are used. K-means segmentation algorithm is used with k = 8, a median filter is applied again with a kernel of (5*5). Finally, threshold and Canny Edge detection are used to segment the tumor part in the image.[2]

2.2.2.Segmentation of Breast Dense in Mammography Image:

At the beginning of the segmentation process, a median filter of kernel 5*5 is applied for smoothing, threshold is applied with n = 50, applying connected labels to extract the breast location in the image and remove the artifacts from the image, then the breast original part added from the original image, Average Threshold is applied on the breast part and median filter to obtain the segmented part to classify in the images.

2.2.3. Post-Processing

The maximum, minimum and number of pixels that exceed 250 gray level are calculated from each image. These values are used to check the type of tumor image whether the tumor is a condensed white part as or a gray level part as by checking on each pixel if it is greater than the minimum calculated before and less than the maximum - the minimum then this pixel is converted to be a white pixel.

2.2.4.Classification

1) SVM Support Vector Machine (SVM) for Brain MRI Classification:

SVM is a state-of-the-art classification method introduced in 1992 by Boser, Guyon, and Vapnik . The SVM classifier is widely used in bioinformatics due to its high accuracy, ability to deal with high-dimensional data. The SVM method has been utilized in a wide range of pattern or image recognition in network fields as it has proven itself more efficient. It also has a better performance in cases of large datasets to simplify the problem. Proposed feature extractions from MRI are passed to fit in SVM Model Table:I with a Linear kernel to train on 75% from Brats2017 ,Brats2015 and Normal datasets. It obtained high results in testing on 25% of the previously mentioned dataset with low time consumption during prediction of class tumor or non-tumor in MRI.

2) Convolution Neural Network (CNN) for Lung CT-Scan Classification:

Deep Learning can lead to a future prediction of cancer and classification. High computation complexity for training, but achieve high accuracy . CNN Deep learning can be used for the detection of cancer existence. Deep Learning can generate features from 2D segmented images which is obtained from the Unet . Unet is a 2d CNN architecture used to classify images and compare it with the ground truth as a training to predict ground truth of cases using dice coefficient cost function.

III. RESULT AND DISCUSSION

3.1. Brain Results

The proposed system has achieved an accuracy of 98% in Brain MRI tumor/non-tumor classification using SVM. This classifier applied on BRATS2015, BRATS2017 datasets and a Normal dataset of 100 patient obtained . In Table I represent a comparison between our system result and the related works results represented After classifying the input image in case of a tumor, segmentation process starts to segment original input image as shown in Fig 4 and get the region of interest as shown in Fig. 5. The segmentation of tumor area in MRI, the introduced system has come to acceptable results.

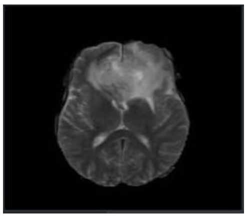

Fig. 1.MRI Brain Original image

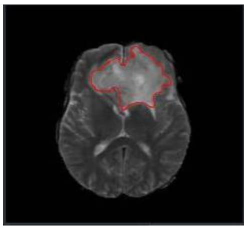

Fig. 2.MRI Brain Segmented image

3.2. Breast Results

The proposed system has achieved and accuracy 95.38% for Breast dense Mammography calculation using ExtraTrees on MIAS dataset. Sample breast input image is shown in Fig. 3

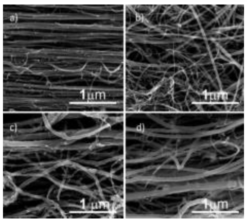

Fig. 3.Breast Mammography Input image

Median, Threshold filters are applied to get the region of interest, classification starts to calculate dense value as shown in Fig.3

3.3. Lung Results

The system also obtained a reasonable accuracy in Lung CT scan cancer detection using CNN (Unet) on LUNA dataset.According to Lung, the input slice of patient is taken to begin process of segmentation which starts with binary threshold as shown in Fig. 4.

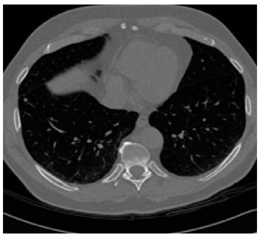

Fig. 4.Lung Ct Scan image

After getting the segmented lungs, region of interest is extracted as shown in Fig. 9 to check out for cancerous nodules.

Fig. 5.Lung Ct Scan Region of interest

V.CONCLUSION

This system comes in handy to doctors as it helps them to classify three types of cancer in different types of images which are MRI, CT scan, and Mammography. The proposed system can classify and segment tumor parts in lung and brain images and calculate breast dense. This system is characterized by the ease of use as it has a friendly interface that allows the doctor to access multiple operations. This system considers achieving a high accuracy in lung cancerous/non-cancerous classification and introducing breast cancer detection and segmentation. Also improving the segmentation of brain tumor. Finally introducing the classification of brain tumor grades (high-grade glioma and low-grade glioma).

ISBN : 978-93-87793-75-0

REFERENCES

[1] S. Sameh, M. A. Azim, A. AbdelRaouf, Narrowed coronary artery de-tection and classification using angiographic scans, in: 12th International Conference on Computer Engineering and Systems (ICCES), Cairo, Egypt, IEEE, 2017, pp. 73–79.

[2] S. Moataz, M. A. Azim, A. AbdelRaouf, Automated fish signals fusion detection for chronic myeloid leukemia diagnosis, in: 26th International Conference on Computer Theory and Applications (ICCTA), Alexandria, Egypt, IEEE, 2016.

[3] M. Clinic, Brain tumor - symptoms and causes, https://www.mayoclinic.org/diseases-conditions/brain-tumor/symptoms-causes/syc-20350084.

[4] Tumor grades and types, abc2.org/guidance/brain-cancer-facts/tumor-grades-and-types.

[5] All About Adult Brain Tumors, https://www.oncolink.org/cancers/brain-tumors/all-about-adult-brain-tumors.

[6] Mammographic image analysis homepage - databases, http://www.mammoimage.org/databases/.

[7] D. L. S. Manisha, Radhakrishnan.B, Tumor Region Extraction using Edge Detection Method in Brain MRI Images, IEEE, 2017.

[8] S. H. Lewis, A. Dong, Detection of Breast Tumor Candidates Using Marker-controlled Watershed Segmentation and Morphological Analy-sis, IEEE, 2012.

[9] K. Greenway, Hounsfield unit — radiology reference article — radiopae-dia.org, https://radiopaedia.org/articles/hounsfield-unit.

.[10] T.S.SubashiniV.RamalingamS.Palanivel, Automated assessment of breast tissue density in digital mammograms, 2010.