A survey on Brain Tumor Classification and Segmentation using MR Images

S.Abirami[1] and Dr.B.Arunkumar[2]

[1]Research Scholar, Department of CSE,Karpagam Academy of Higher Education

[2] Associate Professor, Department of CSE, Karpagam Academy of Higher Education

Coimbatore,Tamilnadu, India.

Abstract: In today's world, image processing plays a vital role in all medical applications. In particular, image segmentation and classification of a brain tumor is one of the most challenging areas. Previously, manual tumor segmentation and classification for MRI images was time consuming and often resulted in errors. After segmentation of tumor it is necessary to classify the tumor either benign or moderated or malignant stage. Now, automatic segmentation of brain tumors focuses on achieving more accurate results. This work analyzes various classification and segmentation methodologies.

Keywords: Feature Extraction, Support Vector Machine, Wavelet Transform, Machine Learning, Gaussian filter

I. Introduction

In the last few days, the researchers have been working towards automatic separation of brain tumors using deep learning techniques. Healthcare is a multifaceted system established with the common goal of preventing, diagnosing, and treating health-related issues or disadvantages in humans. Detecting a brain tumor is a difficult task in the field of medical imaging. This work suggested a particular classification of organized brain imaging and different techniques.

Brain tumor segmentation is involved in various processes such as analyzing, reducing and classifying the tumor. Thus, the separating of the text is a time-consuming process and by improving the automatic separation methods, to give a fair share, it has become an attractive and attractive search function in recent days. The current level of component acquired by the learning methodology to make them better competes in performing this work.

The dislocation of the scalp is highly dependent on the perception of the brain or the leadership of large cells and will remove uncoordinated and irregular cells. The cell group affects the normal functioning of the brain and destroys good cells. The human brain's normal cells are composed of White Matter, Gray Matter, and CSF. Typically, timely diagnosis of brain tumors at the right time is important to add value to improve the course of treatment [2].It is important to find the exact segmentation of the tumor area in medical diagnosis and treatment planning. Segmentation of the tumor in the brain is difficult and can sometimes lead to errors. Therefore, computer-based diagnostics are needed to achieve accurate results. This study consists of a survey of MR imaging techniques and processing methods in-depth study for automatic segmentation and classification of brain MRI.P.Shantha Kumar et.al focusing on various brain tumor segmentation techniques[23].Usually includes brain image processing,

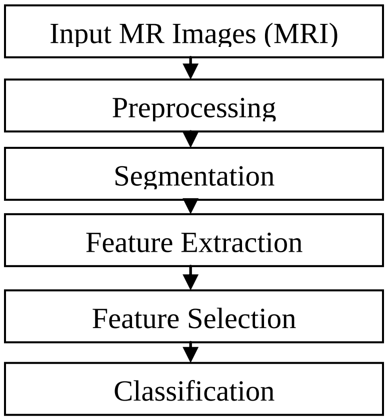

Figure 1. Steps involved in Classification of MR Images

II.Brain Imaging Techniques

Brain research allows scientists and clinicians to predict the problem with the human brain. To identify tumors and their similarity, several advanced CT and MRI scans are available. They play an important role in the diagnosis and treatment planning. Each process guide shows different information from the diagram.

A. Computed Tomography (Ct)

They are invasive and non-contrast imaging techniques to identify tumors. It builds images by absorbing X-rays by network. X-ray images depend on the degree of X-ray absorption. Bone tissue and tissue that absorbs X-rays are the same as air, water, soft tissue. Compared with MRI, CT is better at detecting brain disorders. So CT images are not widely used in tumor studies.

B Magnetic Resonance Image (MRI)

It is a non-invasive imaging technique. It provides anatomical and structural information about the brain. Due to its non-invasive nature, MRI is mostly used for brain tumor analysis. Conventional MR imaging includes standard T1-weighted, T2-weighted, and fluid-attenuated inversion (FLAIR) imaging [11]. In T1-weighted images, brain tumors are hypo or izointense, and in T2-weighted image tumors, they are hyper or izointense. Images with T1 weight have low signal for edema, inflammation, water content, tumor, high signal for fat. Images with weight T2 have the opposite signal for T1. In some cases it is used to improve visualization and adjust the tumor contrast border. It is achieved by means of contrast agents that contain substances such as gadolinium chelates.

Increases the contrast of different images on a different level. It results in stronger T1 and weaker T2 signal. So most often the enhanced T1 is used. Weighted images 1 provide structural information about the brain that can be used to identify healthy tissues, if the area of interest to us is healthy tissue, and then we can use a T1 image. In contrast to T1 the enhanced image boundary of the tumor appears to be brighter which can be used to identify active lesions in the brain [11]. In a T2 weighted image a tumor and edema appear brighter so if the research is on segmentation T2 weighted images are very useful. The Flair image can be used to separate edema from CSF because the watermark is suppressed in this imaging. Today much research focuses on identifying healthy and pathological tissues for further treatment. In that case, a combination of T1, T2, T1 images enhanced Contrast, Flair for accurate results.

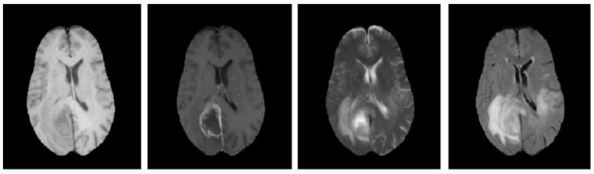

Figure 2.Contains sample image of T1 image, contrast enhanced T1 image, T2 image, FLAIR image.

III.Preprocessing

The primary task of preprocessing is to enhance the quality of the MR image and to make it suitable for further processing by a clinical specialist or machine vision system. Helps to improve the sound ratio from pre-processing signals and enhance the visual appearance of MR images.

A.Noise Removal

Noise appears in images from various sources. It degrades the performance of image processing algorithms. The main purpose of this technique is to store the edges in the image while removing noise. Therefore, non-linear filters and transformation filters are mainly used to remove noise while storing information in the image while removing noise.

Among these filters, the transformation-based filter performs well. Ayse Demirhan et al. the anisotropic diffusion filter used to remove the noise in this inner part of the region are smoothed and the edges are preserved. Anitha et al. Uses an unfiltered middle filter to discolor the image.

B.Intensity Non-uniformity

In solid MR signals from a single tissue vary due to the unity of the RF fiber, the sound wave. These images appear to be identical but this change results in misalignment and segmentation [4]. Powder formulations and modified systems are often used for Bias correction.

C.Skull-stripping

The process of removing the skin / muscle, scalp, skull, Dura, and cerebro-spinal fluid (CSF) from MR images is called skull stripping. The quality of diagnosis and processing can affect if unnecessary tissue is present in the MRI. Therefore, one of the important steps in brain imaging is skull stripping where the brain tissue is completely segmented from the skull[18]. Many skull stripping methods have been developed on the basis of intensity, morphological method or hybrid methods. Some tools also available for skull stripping are Brain Extraction Tool1 (BET), Brain Extraction Based on nonlocal Segmentation Technique (BEaST), Robust Learning-Based Brain Extraction5 (ROBEX), Brain Surface Extractor4 (BSE).

Asit Subudhi et al. Used histogram-based technique followed by morphological surgery to strip the skull. Shaswati Roy used adaptive intensity thresholds followed by morphological operations. Rosniza Roslan et al proposed an applied Otsu threshold and morphological approach to stripping the skull.

D.Image Registration

The conversion of different data sets into a single coordinate system is called image registration. This is an important step in medical image processing when clinical application requires additional information from different images [8]. Many lesions are visible in the MRI image sequence, such as T2 weighted and FLAIR, but their spatial and anatomical resolution are poor, but in T1 weighted images, the lesions are not visible and are high resolution. So by registering these images we get additional information from all the images [8].

E.Intensity Normalization

Images of the same person obtained with the same imaging sequence at different times on the same scanner or different scanners have different intensities. Also, automatic segmentation and classification creates problems when using strength-based algorithms. So, to prevent this, intensity normalization is done in preprocessing. In most cases, histogram-based techniques are used for intensity normalization.

IV.Feature Extraction

Classification is usually done on the basis of discriminate functions. The extraction of functions starts from the initial position of the measured data and the constructed derived values, which are to be informative and non-redundant, which facilitates the subsequent steps of learning and generalization and in some cases leads to better human interpretation. Element extraction is related to dimensional reduction. When the input data is too large to process and contains redundant data, it is transformed into a reduced set of functions. Element extraction is used to extract important information from the raw image that is useful for classification and to express it in a low dimension. The main goal of function extraction is to maximize similarity within a class and minimize similarity between classes. The element extractor generates a vector of elements, which is given as input to the classifier. The accuracy and degree of convergence of element extraction methods have a major impact on the classification results.Some of features and their extraction techniques are given below,

A.Spatial Features

These characteristics are based on the level of gray, amplitude, spatial dispersion. An image histogram can be used to extract these characteristics.

B.Edge and boundary Feature

Edges are used for estimating object boundaries. If the edges are accurately identified, the finely placed object and its characteristics such as perimeter, area, and shape will be easily identified. The information in the form of an image is contained in the borders. Therefore, once we find the edges, it is easier to segment a given region, improve the region, or classify them. But it is a difficult task in image processing.

i)Robert technique

Operator Robert Cross performs a simple, fast 2D gradient measurement on the image. Highlights the region of high spatial frequency that is nothing but edges, it uses 2x2 convolution cores. They are very sensitive to run. They show weak responses to genuine edges.

ii)Sobel technique

It works better compared to the Robert method. At each point in an input grayscale image, it finds the approximate absolute gradient size. It uses two 3x3 convolution cores to identify horizontal and vertical edges from the image. Sobel is slower compared to Robert, but their large cores even out the input image evenly, making them less sensitive to noise. Compared to Robert, the high output frequencies give similar edges. Due to its smoothing, output results in thinning of the edges.

iii) Canny technique

It is an optimal edge detector works in a multi-stage process. First, the image is smoothed using a Gaussian filter, and then a 2-D first-order gradient derivative operator is applied to highlight the areas of the image. This procedure defines the problem associated with edge discontinuities by identifying strong edges and they retain weak edges. The gradient image of this technique contains a ridge at the edges, so this algorithm tracks the ridges and sets zero to all pixels that are not actually at the top of the ridge. This procedure is known as non-maximal suppression. This monitoring procedure is achieved by means of a threshold. The effect of the canny method depends on the width of the Gaussian core, the threshold values. Although a powerful edge detector finds a discontinuity of intensity in images, it does not guarantee that everyone matches the edges.

C.Shape based feature

This refers to the physical structure covering the area, perimeter, circle, and so on. Shape-based features are mostly used for object identification, finding and matching shapes. The shape of the object is obtained from the extraction of external borders based on color, content.

D.Intensity based features

This includes the mean, variance, median, standard deviation, skewness. In order to achieve the statistical method first to be based on design features. In this way, the properties of the individual pixel elements measured spaces interaction. According histogram based technique can be used to derive design features.

E.Texture Features

The torsion feature repeats the pattern of information between regular intervals. It represents a low-level feature in an image, it can be used to describe the content of the image. The characteristic of the entropic theory includes contrast, relation, energy, inertia, entropy of difference, difference of difference, sum difference. The extraction of the image feature plays a key role in medical image attachment applications.

i) Model based Texture Feature Extraction

One of the most widely used sample textured models is the Markovian model. It finds the conditional probability of a given pixel based on the intensity of the neighbor. Random Field Models, Auto Register Models There are other base models that are used to add features.

ii) Statistical based Feature Extraction Methods

Statistical methods characterize the texture based on a non-deterministic property of gray image levels. Some frequently used statistical methods.

a) Gray Level Co-Occurrence Matrix(GLCM)

GLCM is a popular method and is a statistical method. It uses the gray co matrix function to generate a gray level co-occurrence matrix or a spatially dependent gray level matrix. It calculates a pixel with an intensity value of x, how often it occurs in spatial relation to a pixel with a value of y. The spatial aspect ratio is the default pixel and the pixel to its right. Each element (x, y) in the obtained GLCM contains the sum of the number of pixels with the value of x in a certain spatial ratio to the value of the pixel y. The number of rows and columns in the GLCM matrix is based on the number of gray levels in the figure. Texture functions use the contents of the matrix to measure the intensity variation in the pixel we are interested in. GLCM has 14 functions.

b) Local Binary Pattern Features (LBP)

LBP is used to find the connection between the pixel and its neighbors. The results of LBP are encoded in a binary word, so it is called as a local binary pattern. This enables the detection of patterns / functions, while at the same time being immune to contrast changes. It marks the pixels of the image by setting the boundary of each pixel and gives binary result. it is a unifying approach based on statistical and structural model.LBP benefits from lower computational costs. The basic principle is to use 3x3 LBP spatial filters. For each pixel in an image, a square is selected around the current pixel, and then the LBP value of the pixel is calculated using the square. After calculating the LBP value for the current pixel, the update is made at the corresponding pixel location in the LBP mask. The corresponding bit in the binary array is set to 1 if the current pixel value is greater than or equal to the nearby pixel value; Otherwise, if the current pixel value is less than the nearby pixel value, the corresponding bit in the binary array is set to 0.

F.Transform Features

Image modification provides information in frequency store. Some of the transition modes like Fourier, wavelet transitions can be used to take the shape of an image from an image. High frequency segments are used for finding edges and boundaries. Angular slits can be used for visual acuity.

i) Discrete wavelet Transform

DWT can be used for signal analysis, denoising, pattern recognition. It is used to hierarchically divide the image. This is useful for processing non-stationary signals. Waves of different frequencies and limited duration are used for conversion. Wavelet transformation provides both image frequency and spatial description. Time information is stored in this transformation process if it is not possible in a normal Fourier transform. Waves are formed by translating and extending a fixed function called the mother wave. DWT uses filtering to divide the decimation signals at the n level to obtain detailed coefficients. The functions are then derived from the coefficients. Mohamed Khalil et al. k- Classify the brain image as normal and abnormal using the nearest neighbor classification. In this article, they use gray-level co-occurrence model (GLCM), local binary pattern (LBP) and histogram of Oriented Gradient (HOG) feature extraction methods. They send feature vectors to each system and compare their effects. Finally, they reported that GLCM has an average accuracy of 98.49, LBP 94.46 and HOG 100% accuracy.

V.Feature Selection

Due to the increase in training time and diagnostic time, sometimes the number of features is very high. In this case, the characteristics that are really important and class discrimination are selected for classification. This is called feature selection. Features give better results than a full set of subset features. This improves the accuracy of the ranking algorithm, reducing the more appropriate. Here are some of the methods used for feature selection.

A. Filter Methods

In this predictive effect, each function is evaluated. Statistical tests are performed to determine the predictive effect for each variable. One way is to find the correlation with the goal. Function with highest correlation is selected. The disadvantage of filter methods is that they consider one variable at a time. Some of the filter methods are

i) LDA

Linear component analysis is used to find the linear combination of properties that characterize two or more classes of a categorical variable.

ii) ANOVA

Variation analysis is similar to LDA. But it worked by using one or more independent features and one consistent feature.

iii) Chi square

It is a calculation measure that is used in a group of experimental groups to assess the feasibility of communication or interaction between them.

B.Wrapper Methods

The wrapping method uses a combination of variables to find the estimated power. It finds the best combination variable. This method test each property compared to the test model it creates with them to evaluate the results. It is not recommended that this method be not applicable to high quality attributes. They are intensive computing.

i) Forward Selection

It uses iterative method, where we start with empty functions then, in each iteration functions are selected that improves the model. When we understand that there is no improvement in the classification by adding function, we stop adding functions.

ii) Backward Elimination

In this method, we start with all the features and remove the least significant feature on each iteration, which improves the performance of the classification, the procedure is repeated until we find an improvement.

iii) Recursive Feature Elimination

It is a greedy optimization algorithm that aims to find the best subset. In repetition, it finds the best or worst function. It then creates a pattern and sets it aside. It then creates the next model with the remaining features until all the features are gone. It then classifies the features based on the order of their elimination.

C.Embedded Method

Embedded methods use built-in methods. I haven't built a model for functional testing. You perform feature selection as part of the model building process. The most popular examples of these methods are LASSO and RIDGE regression. They have a built-in penalty feature to reduce overfitting.

VI.Machine Learning and Deep Learning Based Algorithms for Segmentation and Classification

A. Machine Learning

Machine learning allows you to learn from examples and experiences without being explicitly programmed. It's a subfield of artificial intelligence. The machine learning algorithm allows computers to train data and search for output using statistical analysis. Using machine learning, computers create models from data so they can make automatic decisions. Two of the most widely used machine learning methods are supervised learning, unsupervised learning.

i) Supervised learning

With this method, the computer has examples of inputs that are marked with the desired outputs. In this method, algorithms are "learned" by comparing the actual output with the desired outputs in order to find errors and modify the model accordingly.

a)Neural Network Classifier

Neural networks are suitable for detecting nonlinear patterns when there is no one-to-one relationship between input and output. It is a method of controlled learning. An artificial neural network is organized into layers of interconnected units or nodes called artificial neurons. If the desired result is already known, it is said that the neural network learns to be controlled, otherwise without control.

Input Layer: Views the data. The amount of neurons in a storage depot is equal to the number of objects in the data.

Output Layer: The number of neurons in the output layer is subject to no classification label.

Hidden Layer: In ANN between input and output layer there are one or more layers called hidden layers. The nodes in the hidden layers apply transformation to the input.

ANN neurons apply an activation function that combines input values, associated weight, and weights and maps the result to an output. After submitting all records, the process is repeated many times. Slow convergence to propagate the return. The disadvantage of ANN is its tendency to overload and overload.

Deepak Ranjan Nayak et al proposed a probable neural network approach to automatically detect pathological brain. A probable neural network is based on Bayesian theorem. In contrast, a limited adaptive histogram balancing technique is used to enhance the image. The Region of Interest is then segmented using a pulse coupled simplified neural network. Features are then generated using a second-generation curvature transformation known as discrete quadrangular rapid transformation (FDCT). This procedure is simpler, faster and less redundant. It then uses the combination of PCA and LDA to select a feature. Eventually a Probable Neurological Network is used for classification and 99.57% accuracy to detect pathological brain.

Varuna Sri et al proposed a classification of a brain tumor based on a probable neural network. First, the image is pre-processed to improve the signal-to-noise ratio. Then segmentation is made based on the Region Cultivation Technique based on the intensity and mean value of the region. Morphological operations are then performed to extract the border regions. DWT and GLCM are used for feature extraction. Finally, the classification of the tumor into normal and abnormal is done through a probable neural network and they have achieved 100% accuracy.

b) Random forest Classifier

Random Forest Classifier is a set algorithm.The combined algorithm means that it combines more than one algorithm of the same or different to classify objects. RF is a multidirectional classifier.It creates a set of decision trees from a randomly selected subset of training. The final class of the test object is determined by adding the votes of different decision trees. The advantage of RF is the ease of training and testing.

Khalid Usman, Kashif Rajpoot supported the RF-based classification of multimodal brain image. They used BraTS data collection images T1, T2, T1C and FLAIR. Pretreatment involves recording (with a T1 image), removing the skulls (by experts), fitting the histogram (to improve the image), finding a confining box (around the tumor). Its novelty is the extraction of working-wave functions from a multimodal image. Functions are separated using the 3D Wavelet Transform. It separates intensity, intensity differences, neighborhood information, and wave-based texture functions from the function image. First, they used KNN to create a tag for classification. The RF classifier then trains several trees from the training vector. When training several trees, the correlation increases and the scattering between the trees decreases. To follow the RF, AdaBoostM2 (Boosting Algorithm) is used for boosting, which connects many weak learners to a powerful algorithm and weighs the specimens instead of retrying them (in bags). Machine learning methods do not give an effective classification result for skewed data, so they used RusBoost to solve the problem by combining sampling and amplification. They achieved a score of 75% in the tumor area and 95% in the tumor area, which is greater than the overlap of the dice reported in the MICCAI BraTS test and takes about 2 minutes to test a new patient.

c) Support Vector Machine

Support Vector Machine (SVM) is a supervised machine learning technique. It can be used for both classification and regression problems. In this, each data item is plotted as a point in n-dimensional space, where n is the number of elements. Then the classification is performed by finding the hyper rovine, which distinguishes these two classes very well. An easy way to describe SVM is a binary classifier. It can separate two classes of data using the hyper plan with the largest margin. Formally in mathematical language, SVMs construct linear separation hyper plans in high-dimensional vector spaces. Optimal classification is achieved when the hyper plan provides the maximum distance to the nearest training data points. If the points are well separated, the classification between the two groups is much clearer [16].

For optimal classification, the decision limit should be as far away as possible from the data in both classes. If the data is linear, a hyper-separation plane can be used to divide the data. Sometimes data sets are inseparable, which means they are not linear. In this case, the core-based SVM is used to solve nonlinear problems. A kernel is a similarity function. It takes two inputs and spits out the degree of similarity they use with a technique known as the core trick, which can be much more flexible. The SVM is an effective tool in high-dimensional spaces. Class separation is often not linear. The application of cores allows making flexible the limits of the decision and leading to a greater performance of classification. It is sensitive to noise a small number of mislabeled examples can significantly decrease performance.

Nilesh Bhaskarrao Bahadure and co-workers used biologically inspired BWT and SVM to detect the brain. They used T1, T2, Flair MRI images from real time. In preprocessing, adaptive improvement is performed based on a modified sigma function, and the threshold technique is used to remove the skull from the MRI image. Segmentation is performed with BWT after spread and erosion operations are applied for efficient tumor segmentation. The texture characteristics are then extracted by the GLCM method. Finally, the classification of normal and abnormal is done with SVM based on the Gaussian core. The experimental results achieved an accuracy of 97.51%, a specificity of 95.2% and a sensitivity of 98.72%. The experimental results also received an average cube score of 0.91.

d) Logistic Regression

A binary classification technique is one that relates to supervised learning based on logistical function, which calculates the weighted sum of the input variable. It involves a probabilistic view of classification. LR is a predictive modeling algorithm that is used when a particular variable is in the binary category. Regression is because we adapt the linear model to the appearance of space. In this method, data sets are analyzed using statistical methods that measure the relationship between the dependent categorical variable and one or more independent variables using the logistic function by calculating the probabilities. The class assignment is then done by applying the probability threshold. The advantage of LR is that it is easy to train, it is very fast in classifying data sets, it is resistant to over-adaptation. The disadvantage is the linear limit of the decision.

e) Naive Bayes Classifier

This is a supervised machine learning technique based on the Bayes theorem. The Naive Bayes classifier is based on the assumption that the characteristics of a class are independent of all other properties. Two types of quantities must be calculated from the data set for the naive Bayes model, these are class probabilities, conditional probabilities. The prediction is then made using the Naive Bayes theorem.

Nidhi gupta et.al suggests classifying brain tumors according to salivator Naïve Bayes. T1W, T2W, FLAIR, T1C image modes are used from JMCD and BRATS data. The development is carried out by Stochastic Dynamic Resonance and Anisotropic Diffusion. Combining different pairs of images, enemy division is used. It captures the image of two groups the first group being T1W, T1C and the second group also T2W, FLAIR. An adaptive and gluteal valve is then used to dissect the tumor. Feature output is done with the gray run length matrix (GRLM). It is a combination of LBP and GLCM. The tumor is then classified into low-grade and high-grade glioma by classifier Naïve Bayes. They compared the result of their classification with SVM, KNN. Distributor Naïve Bayes scored 9.47 for BRATS 97.86 for the JMCD which is high according to the above guidelines. They also take 3.18 and 5.06 s for BRATS and JMCD Time which is also lower than SVM, KNN.

B.Unsupervised learning

Sometimes the data is not labeled in this case unsupervised learning is particularly valuable. It tries to find common ground among its input data .Hidden patterns in a data set are detected by unsupervised learning .can also have a feature learning goal. Through this method the computing machine automatically detects the representations that are needed to classify raw data. Unsupervised learning is also referred to as Grouping Technique because it forms clusters in the dataset to classify them.

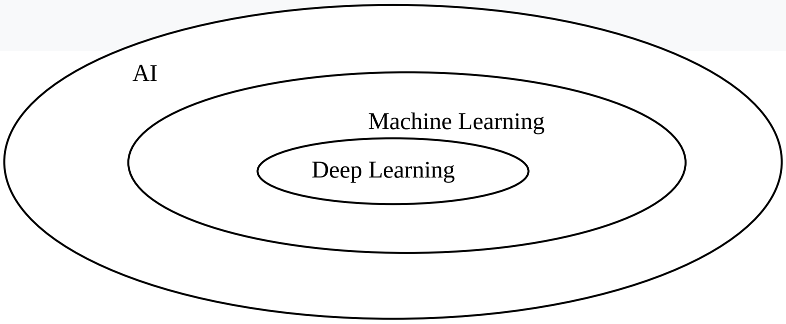

Figure 3.Differentiation of AI, Machine Learning, Deep Learning

Some of machine learning algorithms are reviewed here,

a) K-Means Clustering

This is the simplest unsupervised algorithm. This algorithm follows a simple method of classifying data given by the number of clusters. It first defines the K centers for each cluster. The data is then taken one by one and added to the cluster based on the nearest cluster center. When there is no waiting point, the grouping is done. In the next step k the centers are recalculated from the previous step. New data links and centers are being made again. And this process is repeated until the centers are no longer moving. This method is fast, robust and easier to understand. In this method the best results are obtained when the data set is distinct or well separated from each other. Not suitable for the dataset that contains overlapping data.

Anitha et al used a K-based method to classify brain tumor. First, magnetic resonance imaging of the brain is preprocessed using a non-local medium filter. To elude in the K-pillar algorithm means that all centroid are very discreetly between them. They also applied the mean-based mean center instead of the mean in the K means. The functions are then extracted using DWT. During classification, the characteristics are first trained with SOM and the resulting filter factor is trained with the neighbor closest to K. Then, the tests are also done at two levels. They compare their results with SVM and the proposed method provides better results. They achieved a maximum accuracy of 96.7%.

b) Self Organizing Map

It is clustering based technique. It is unsupervised models of neural networks in which the network layer to form their own classifications of training without external help.SOM consists of two layers of nodes. In SOM the input node directly connected to output layer without any hidden layer. The nodes in the input layer denotes features and the output layer also known as “kohenen layer” or SOM Layer represent low dimensional visualization of data. Output nodes are arranged in the form of topological architecture number of nodes in the output layer denotes maximum number of clusters. Typically the network topology is arranged in either rectangular or hexagonal grid. Initially all the weights are initialized with random variables. For each input pattern, the neurons in the output layer compute respective value of discriminate function the neuron with smallest value of discriminate function is declared as winner Best Matching Unit (BMU) .Mostly Euclidean distance is used as discriminate function. Then winning neuron determines the spatial locations of neurons whose radius are within the radius of BMU (closest node) then the weight of BMU and closest nodes are altered to make them more like input vector. This process is repeated for all records. They have advantages like learn from data, fault tolerance. It always searches for optimal solution.

The disadvantage of SOM is that it does not build a generative data model, slow training, difficult to train against data that is slowly evolving.

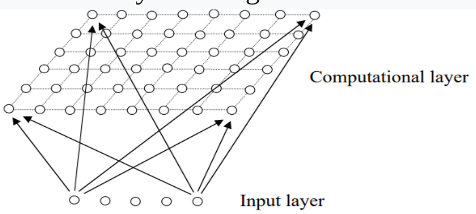

Figure 4.Architecture of SOM

Ayse Demirhan et al. the proposed SOM-based human brain segmentation. SOM is one of the unsupervised learning algorithms. NOM is learned from input data and has more output neurons than usually specified. So the output of SOM must be grouped. Typically an additional NN is used in the output layer to group the output. In this article the image is pre-processed to remove noise by means of an anisotrophic diffusion filter, which has the advantage of sharpening the edges of the object while removing noise. Then the images are recorded. After that, the skull is stripped using the posture threshold and morphology operation. Mostly the extraction of Brain Features is done with DWT. The drawback to using DWT is the Translation Variant transform. It brings different features for two identical images. So they used SWT (stationary wavelet Transform) to overcome the problem specified above. And they used entropy; absolute mean deviation, standard deviation as a characteristic texture for classification. After that, the SOM unsupervised learning algorithm is used and no additional NN is used in this methodology to group the output. Instead, they developed an algorithm based on the success of the best tissue segmentation neuron in the histogram, using the region of the manually labeled image. They then used the LVO algorithm, which uses tagged data, to fine-tune the weight vector of the SOM. Eventually, the brain region becomes white matter, gray matter, CF, tumor, edema. They give a result of 91.02% for WM, 88% for GM, 96.1% for CSF, 61.01% for Tumor, 77.02% for edema.

C.Deep Learning

Artificial Neural Network Deep Learning is a sub-area of machine learning. The deep learning framework extends traditional neural networks (LNs) by adding more hidden layers to the artificial neural network architecture between the input and output layers. DL can model more complex and nonlinear relationships. This concept has aroused the interest of researchers in recent years due to its good results, thanks to which it has become the best solution to many problems related to the analysis of medical images, natural language processing etc.A convolutional neural network is a deep learning architecture mainly used for image classification are the most popular deep neural networks constructed of neurons with teachable weight and prejudice. Each neuron takes several inputs, takes a weighted sum from them, and passes it to the activation function and responds with outputs. Most CNN architectures include a convolution layer, a pool layer, and fully interconnected layers. Input to CNN is width x height x depth matrix, here depth is R, G, B channels.

Convolution layer

CNN contains one or more convolutional layers. The convolution layer consists of a set of filters that can be learned. Each filter is considered a feature identifier. It applies a convolution operation to the input. The most important parameter in a convolutional neuron is the size of the input. The output of the convolution layer is the activation map or feature map.

ReLU Layer

A ReLU Layer is usually placed after the Convolution Layer. The purpose of this layer is to introduce nonlinearity into the system. This layer changes all negative activations to zero.

Pooling Layer

The pool layer is used mostly after the solution layer. The function of the pool arrangement is to reduce the representation size of the representation size. By reducing the space size, it reduces the parameters there, thereby reducing the calculation and fit. Every layer of the merge layer works independently, and MAX operations are used to resize it remotely.

Fully Connected Layer

All neurons in the fully connected layer are connected to all activated neurons in the previous layer. This layer takes an input volume and produces an N-dimensional vector where N is the number of classes. The output of this layer is calculated by matrix multiplication followed by bias offset.

Data Set

BRATS, OASISS, NBTR, and Brain Webs are some of the datasets available for brain tumor studies. Tumor Sim is simulation software used to generate images of tumors from healthy images.

VII. Conclusion

Automatic segmentation of the brain tumor is extremely stressful. The product of the medical image is growing rapidly in the field and plays an important role in initiating diagnostic problems. More recently the development of imaging techniques and machine learning tools allows researchers to initiate and visualize a tumor for other treatments. This analysis focused on the whole process involved in automatic segmentation and tumor classification. In this paper the various imaging techniques found in MRI are discussed. This work is designed to provide a comprehensive overview of the processes that occur in brain MRI and the function of machine learning and depth of learning algorithms and the correction of their automatic part of the tumor from brain MRI.

VIII.References

1. P. Thirumurugan, P.Shanthakumar,"Brain tumor detection and diagnosis using ANFIS classifier", in International Journal of Imaging Systems and Technology, Wiley Blackwell, Volume: 25,Issue:4, (2016),pp. 157-162.

2.Nidhi Guptaa , Pushpraj Bhatelec, Pritee Khannaba ," Identification of Gliomas from brain MRI through adaptive segmentation and run length of centralized patterns ",in Journal of Computational Science,vol25,(2018),pp.213-220.

3. Asit Subudhi, Jitendra Jena and Sukanta Sabut, "Extraction of Brain from MRI Images by Skull Stripping using Histogram Partitioning with Maximum Entropy Divergence", in IEEE International Conference on Communication and Signal Processing (ICCSP),2016.

4. John G. Sled, Alex P. Zijdenbos and Alan C. Evans" A Nonparametric Method for Automatic Correction of Intensity Non uniformity in MRI Data", in IEEE Transaction on Medical Imaging,(1998),pp.87-97.

5.P.Shanthakumar,P.Ganesh Kumar" Performance Analysis of Brain Tumor Diagnosis Based on Soft Computing Techniques", in American Journal of Applied Sciences, Science Publication, Volume: 11, Issue: 2,(2014), pp.329-336.

6.K.Batri and D. Sundhara Raja, P. Thirumurugan, D. Ramkumar," Automated detection of glioblastoma tumor in brain magnetic imaging using ANFIS classifier",in International Journal of Imaging Systems and Technology, WILEY BLACKWELL,Volume:26,Issue:2,(2016),pp. 151-156.

7. Ayes Demirhan ,Mustafa Toru and Inan Guler, "Segmentation of Tumor and Edema Along With Healthy Tissues of Brain Using Wavelets and Neural Networks", in IEEE journal of Biomedical and health Informatics,vol 19,(2015),pp.1451-1458.

8.Jue Wu and Albert C.S. Chung," Multimodal Brain Image Registration Based on Wavelet Transform Using SAD and MI", in Medical Imaging and Augmented Reality, Second International Workshop,(2004)Aug 19-20.

9.Mohammed Khalil,Abib Ayad and Abdellah Adib,"Brain image classification system evaluation of feature extraction techniques in MR" in Procedia Computer Science,vol 127,(2018),pp.218-225.

11.Stefan Bauer1, Roland Wiest, Lutz-P Nolte and Mauricio Reyes," A survey of MRI-based medical image analysis for brain tumor studies", in Physics in Medicine and Biology, Jul 7;58(13),(2013),pp.97-129.

12.P.Shanthakumar,P.GaneshKumar," Computer aided brain tumor detection system using watershed segmentation techniques", in International Journal of Imaging Systems and Technology, Volume: 25, Issue: 4,(2015),pp. 297-301.

13.Jasmine Hephzipah Dr.P.Thirumurugan," International Journal of Pure and Applied Mathematics", Special Issue on Recent Research in Science, IJPAM,Scopus, Volume: 116,Issue: 17,(2018),pp.1669-1676.

14. Nilesh Bahadure and Arun Ray, "Image Analysis for MRI Based Brain Tumor Detection and Feature Extraction Using Biologically Inspired BWT and SVM", in Hindawi International Journal of Biomedical Imaging, Vol(2017), pp 1-12.

15.SonuSuhag, Lalit Mohan Saini,"Automatic Brain Tumor Detection And Classification Using SVM Classifier", in International Journal of Advances in Science Engineering and Technology, vol- 3, (2015) Issue-4.

16.P.Thirumurugan and P. Shanthakumar N. Ezhilmathi," An Automated Brain Tumor Detection and Segmentation Methods Using MRI Images- A Review", Journal of Advanced Research in Dynamical and Control Systems, IASR, Volume: 1, Issue: Special Issue, (2018),pp. 1678-1681.

17.Shaswati Roy and Pradipta Maji,"A Simple Skull Stripping Algorithm for Brain MRI",in IEEE Eighth International Conference on Advances in Pattern Recognition(2015).

18.Javeria Amina, Muhammad Sharifa, Mussarat Yasmin, Steven Lawrence Fernandesb,"Big data analysis for brain tumor detection: Deep convolution neural networks", in Future Generation Computer Systems, (2018).

19.Yasmin, Mussarat, Sharif, Muhammad, Mohsin,Sajja,Azam and Faisal,"Pathological Brain Image Segmentation and Classification: A Survey in "Current Medical Imaging Reviews", vol 10, Number 3,(2014), pp. 163-177.

20.Jin Liu, Min Li, Jianxin Wang, Fangxiang Wu, Tianming Liu and Yipan," A survey of MRI-based brain tumor segmentation methods" in Tsinghua Science and Technology ,Volume: 19, Issue: 6,(2014),pp.578 – 595 .

21.P.Shanthakumar, P.Ganeshkumar," Performance analysis of classifier for brain tumor detection and diagnosis" in Computers and Electrical Engineering, Elsevier Science, Volume: 45, (2015),pp.302-311.

22.N.VarunaShree,T.N.R.Kumar,"Identification and classification of brain tumor MRI images with feature extraction using DWT and probabilistic neural network" in Brain Informatics, (2015),pp.23–30.

23P.Shanthakumar,P.Thirumurugan, N.Lavanya Devi,"A survey on Soft Computing Based Brain Tumor Detection Techniques using MR images", in Journal of Seybold Report,(2020),Vol 15,Issue 9,pp.2951-2962.