وزارة التعليم العالي والبحث العلمي

جامعة بابل/كلية العلوم للبنات

قسم علوم الحاسبات

Minimum Distance Classifier for MRI Images based on and Histogram's features

مشروع تخرج إلى قسم علوم الحاسوب في كلية العلوم للبنات/جامعة بابل كجزء من متطلبات نيل شهادة البكالوريوس في علوم الحاسبات

اعداد الطالبة

آيات عبد الكاظم فزع

بأشراف الاستاذة

أشراق عبد الأمير يحيى

بِسْمِ اللَّـهِ الرَّحْمَـٰنِ الرَّحِيمِ

{يَرْفَعِ اللَّهُ الَّذِينَ آمَنُوا مِنْكُمْ وَالَّذِينَ أُوتُوا الْعِلْمَ

دَرَجَات} المجادلة : الاية١١

صدق الله العلي العظيم

الاهداء

إلى من بـــــــها أعلو، وعليها أرتكز، إلى القـــــــلب المعطاء
(والدتي و الدي)
إلى من بذلوا جهدًا في مساعدتي وكانوا خيرَ سندٍ
(إخــواني وأخواتي)
الى مشرفة البحث الست أشراق عبد الأمير
التي ساهمت في نجاح البحث
إلى أسرتي إلى أصدقائي وزملائي ....
إلى كل من ساهم ولو بحرف في حياتي الدراسية.....
إلى كل هؤلاء: أهدي هذا العمل، الذي أسال الله تعالى أن
يتقبله خالصًا....

الشكر والتقدير

نشكر الله العلي القدير الذي أنعم عليَّ بنعمة العقل والدين القائل في محكم التنزيل

أتَقدم باحترام وتقدير كبير إلى جميع الأساتذة والباحثين الذين قدموا معلومات ونصائح قيمة خلال إعداد هذا البحث

نشكر كل من ساهم بأية طريقة في إثراء هذا العمل البحثي وجعله أكثر اكتمالًا وجودة.

Table of contents

Figures

List of Tables

Abstract

Computer techniques play important role in medical fields, especially, in the classification and disease diagnoses. In this work, Brain MRI is classified into normal and abnormal by using minim distance classifier after extracted features by histogram method. The system shows high efficiency through practical experiments where the accuracy of system is reached to 80%.

Chapter One: Overview

1.1 Introduction

In recent decades, and to this day computer technology has been used in applications and various fields including the medical field, which prompted many researchers to employ this technique in the design of decision support systems using many of the algorithms and methods for this purpose.

Brain tumor is one of the major causes for the increase in mortality among children and adults. A tumor is any mass that results from abnormal growths of cells in the brain. It may affect any person at almost any age [1].

Many procedure and diagnostic imaging techniques can be performed for the early detection of any abnormal changes in tissues and organs such as Computed Tomography (CT) scan, Magnetic Resonance Imaging (MRI), X-ray, and Ultrasound. MRI is a primary medical imaging modality because it provides rich information for excellent soft tissue contrast which is especially useful in neurological studies [2].

Texture analysis is an important task in many computer application of computer image analysis for classification, detection or segmentation images. Texture extraction methods can be classified into three major categorys: statistical, structural, and spectral [3].

Data mining seeks a solution for real world health problems in the diagnosis and handling diseases. Several data mining techniques were used by Researchers in the medical field such as decision tree, k-means, fuzzy c-means, k-nn and neural network [3].

1.2 Related works

Rakesh Kumar and Jyotishree(2016) are implemented a Classifier System for extraction of tumor region and diagnosis of brain tumor. The system was used to preprocess of brain MR images in many steps such as, noise removal, skull striping and intensity normalization. The free noise brain MR images used to extract the Texture and intensity features. Then method uses multiclass Support Vector Machine (SVM) to classify five types of tumor based on the WHO grading system i.e Astrocytoma(Grade-I), Glioblastoma Multiforme(Grade-IV), Meningioma(Grade-II), Medulloblastoma (Child tumor) (Grade-IV) and metastatic melanoma(Grade-III)[4]. **Sudipta Roy and Shayak Sadhu(2016)** they supposed a system includes a way characterized that is fast and effective to diagnosis of tumor to identifying the right class of brain tumor. The classifier system based on Many steps, beginning with taken MRI image as input and normalize its. After that, the system reduced a redundancy of data that acts the input of classifier by extracting the features from image. The classifier produced the output using each tuple of extracted features. Fuzzy Inference System (FIS) based classifier known as Adaptive Neuro Fuzzy Inference System (ANFIS) was used to classify input[5]. **Ch.Amulya and G. Prathibha(2016)** they proposed a system includes a classification of images was done by feature extraction that using two methods for extraction which are Scale Invariant Feature Transform (SIFT) and Speeded Up Robust Features (SURF). SIFT method is used to detect the images with larger corners and extract them. SURF, the name itself represents a speed method to extract the features when compared to SIFT. KNN classifier is used to classify the images based on the features extracted from both techniques. So these combined processes are applied to classify tumor and non-tumor images more accurately[6].

1.3 Aim of the research

The aim of this research is to present an automated method that assists diagnosis of normal and abnormal MRI brain, to avoid the human error in manual interpretation of medical image content.

1.4 Research Layout

This research is organized as follows:

Chapter Two: The overall objective of this chapter is to present fundamentals, details, and characteristics of all approaches which have been used in proposed research, where this chapter starts with short introduction to image processing and then it gives an explanation about the methods that have been used in classification.

Chapter Three: This chapter presents the designed steps of the entire system's stages, displays the implementation results of experiments, and a discussion on the obtained results.

Chapter Four: The derived conclusions from the proposed research and some suggestions to enhance the proposed research have been presented in this chapter.

Chapter Two: Theoretical Background

2.1 Introduction to Image Processing

Image processing is always an interesting field as it gives an improved pictorial information for human interpretation and processing of image data for storage, transmission, and representation for machine perception. This field of image processing drastically improved in recent times and extended to various fields of science and technology .In recent times this branch is known with a new name called computer vision[7].

Image processing is processing of images using mathematical operations for which the input is an image, a series of images, or a video, such as a photograph or video frame, the output of image processing may be either an image or a set of characteristics or parameters related to the image[3].

2.2 Digital Image Representation

A digital image is composed of a finite number of elements, each of which has a particular location and values. The elements of a digital image are called picture elements, image elements, pels or pixels (the most widely used to indicate the elements of a digital image are

Pixels [8].

There are three kinds of digital images : Binary , Gray-scale, and Color. The image type that was considered in the present research is Gray-scale. Gray-scale images are referred to as monochrome, or one-color images. They contain brightness information only , without color information. The number of bits used for each pixel determines the number of different brightness levels available. The typical image contains 8 bit / pixel data , which allows us to have 256 (0-255) different brightness (gray) levels[9].

2.3 Pattern Recognition Principles

The patterns refer to those objects that can be perceived. , while pattern class can be defined as a set of patterns that share some properties in common. Pattern recognition is a process of categorizing any sample of measured or observed data as a member of one of the several classes or categories[10].

Pattern recognition system consist of three phases: data acquisition , feature extraction, and decision classification, as shown in figure(2.1).

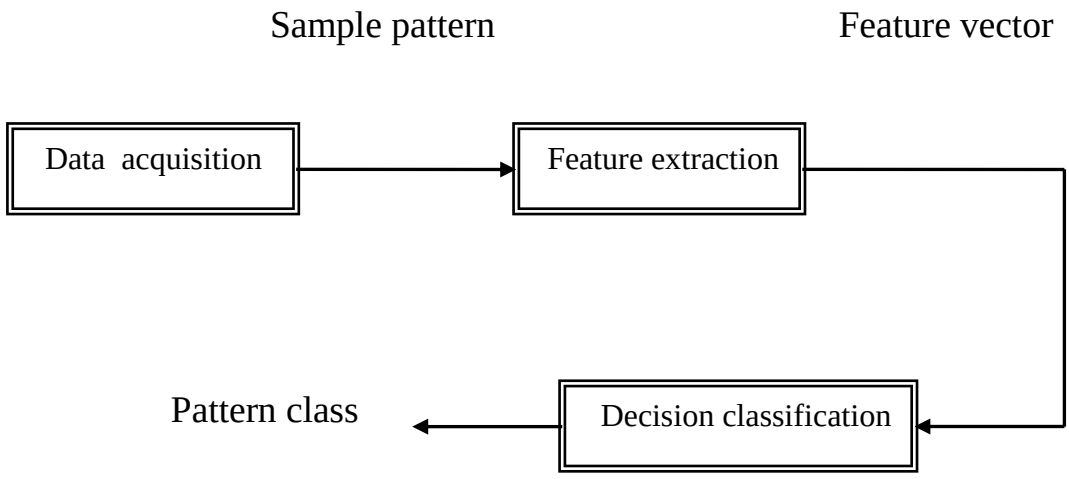


Figure (2.1) Components of a pattern recognition system.

2.4 Feature extraction

Feature extraction is the operation of extracting helpful information from image or identifying characteristics found within the image, these characteristics are used to describe the object. The techniques to measure the features are known as feature extraction techniques, which operate on an image array and produce a feature vector. Patterns can be described by a set of features [7].

In this research, six textural features based on histogram are extracted of every image. The histogram is one of the most first-order texture analysis measures to calculate texture. The main advantage of this approach is its simplicity to characterize the data. The histogram of an image is a function that provides the frequency of occurrence for each intensity level in the image.

Statistical features can be driven from a histogram , that provide us with information about the characteristics of the gray level distribution for the image. The histogram probability P(g) is defined as[3] :

$$P(g) = \frac{N(g)}{M}$$

Where N(g) is the number of pixels with gray-level (g), and M represents the total number of pixels in the image . The histogram features are the mean , standard deviation , skewness , energy , entropy , and Kurtosis.

1- Mean

The mean is the average value that tells us something about the general brightness of the image , the mean is defined as follows :

$$\bar{g} = \sum_{g=0}^{L-1} g \cdot p(g)$$

Where L

represents the total gray-levels , therefore, the gray – level g is from 0 to L-1

2- Standard Deviation

The Standard deviation is the square root of the variance. It provides us with information about the contrast. It is defined as follows :

$$\sigma g = \sum_{g=0}^{L-1} \sqrt{(\bar{g} - g)^2 . p(g)}$$

3 – Skewness

The skewness measures the asymmetry about the mean in the gray-level distribution, it is defined as :

$$\text{Skew} = \frac{1}{\sigma^3 \sum_{g=0}^{L-1} (\bar{g} - g)^3 \cdot p(g)}$$

4 – Energy

The energy is a measure that provides us with information about how the gray-levels are distributed. It is defined as :

$$\text{Energy} = \sum_{g=0}^{L-1} [p(g)]^2$$

5 – Entropy

This measure tells us how many bits we need to code the image data, it is defined as :

$$\text{Entropy} = - \sum_{g=0}^{L-1} \sum p(g) . \log_2 [p(g)]$$

6 – Kurtosis : It defined as

$$\text{Kurtosis} = \frac{1}{\sigma^4 \sum_{g=0}^{L-1} (g - \bar{g})^4 . p(g) - 3}$$

2.5 Supervised and Unsupervised Classification

The classification problem will simply find a separating surface that partitions the known prototypes into correct classes. This separating surface is expected to be able to classify the other unknown patterns if the same criterion is used in the classifier.

Pattern recognition is an integral part of image processing and finds its to document classification, and remote sensing. The goal of pattern recognition is to establish a relationship between a pattern and a class label, a number of pattern classification techniques are used for the recognition of patterns. Pattern recognition techniques are mainly of two types:[11]

1. Classification methods based on supervised learning.
2. Classification methods using unsupervised techniques.

The supervised method for pattern labeling requires the user to select representative training data (samples) for each of a predefined number of classes to train or teach the classifier to determine the decision boundaries in feature space, and such decision boundaries are significantly affected by the properties and the size of the samples used to train the classifier. The minimum distance classifier and the maximum likelihood classifier are some of the frequently used supervised algorithms[8].

Unsupervised pattern recognition methods (self- organized learning) are less dependent on user interaction. Unsupervised classifiers learn the characteristics of each class (and possibly even the number of classes) directly from the input data without the user having foreknowledge of the existence or names of those classes. In unsupervised approach , the machine partitions the entire data set based on some similarity criteria, it is performed most often using clustering methods. This procedure can be used to determine the number and location of the spectral classes into which the data falls[12].

2.6 The Minimum Distance Classifier

An unknown pattern is assigned to the class to which it is closest in terms of a predefined metric. Distance functions are used to measure the similarity or dissimilarity between two classes of patterns. The smaller distance between two classes of patterns is the larger similarity between them[8].

There are different kinds of distance measurement that are used to recognize a pattern or classify the pattern class. The distance measurements are used as a measure of similarity, where any two patterns within the same class are similar (i.e. the distance between vectors corresponding to patterns from the same class are small). There are many of distance functions[12].

The following distance functions generalized a distance measure between two data objects or patterns X and Y .

1. Minkowski distance :

$$d(X,Y) = \sqrt[p]{\sum |X_i - Y_i| p}$$

2. Manhattan distance :

$$d(X,Y) = \sum_{i=1}^{n} |X_i - Y_i|$$

3. Euclidean distance

$$d_2(X,Y) = \sqrt{\sum_{i=1}^{n} (X_i - Y_i)^2}$$

Chapter Three: The proposed System

3.1 The Structure of The proposed System:

The proposed method contains different steps, including the data collection, feature extraction through histogram and classification through small distance technique. The database contains both normal brain and abnormal brain images. In the first, features of the MRI brain images are extracted through histogram. There are various features obtained from the image such as mean , standard deviation , skewness , energy , entropy , and Kurtosis. Classifier technique are used is used for diagnosing MRI brain images into normal and abnormal.

3.1.1 Image Acquisition:

A 100 Brain images are collected from Marjan hospital of Babylon city in Iraq with the class 1 refers to existence of abnormal tumor and class 0 for normal.

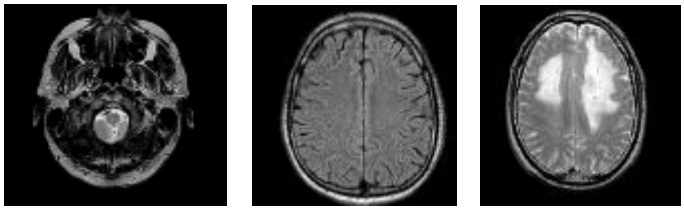


Figure (3.1) : set of MRI Images.

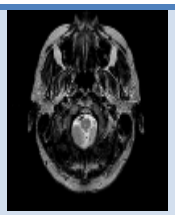
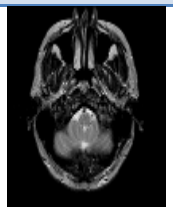
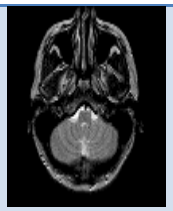
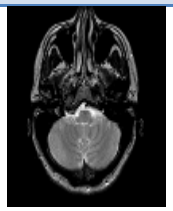
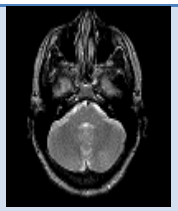
3.1.2 Features Extraction Histogram Technique:

Feature extraction is the process in which, all the features are extracted for accurate classification of MRI brain images. In this research, feature extraction is done using the histogram.

Histogram method uses on the images to obtain 6 features vector.

Table(1) shows part (6 features of 5 images of 100 images) of database resulted from histogram set of brain images. Features are entered to the classifier tool as inputs.

Table (1): Part of extracted features for various images

Input image	Mean	standard deviation	Skewness	energy	entropy	Kurtosis
	28.7829	1.0000	1.6722e+005	0.3100	4.2156	2.2269e+007
	50.8862	1.7260	3.8744e+004	0.2385	4.8111	4.1632e+006
	39.3208	1.0270	1.3859e+005	0.3061	4.1981	2.0564e+007
	37.3953	1.0155	1.7900e+005	0.3175	4.2216	2.5257e+007
	63.0900	1.8006	1.0295e+004	0.1299	5.5211	2.0980e+006

3.1.3 Classifier Technique:

The classification process is done using the tool which are described in the (chapter 2 section 2.6), set of experiments are done and their results explained in the table(2)

Table (2): set of experiments and their results

Size	Performance (Accuracy)
Performance of (80%train-20%test)	75
Performance of (70%train -30%test)	80
Performance of (50%train -50%test)	72

This section deals with clarifying the results of the proposed research, the dataset of features vectors for MRI has been divided into train and test differ in size as mentioned in the table(2) , the performance of the Suggested Classifier System in range of (72 - 80).

Chapter Four: Conclusions and Future work

4.1 Conclusions and future work:

This research is proved its success as a classifier system of brain MRI images. The Minim distance classifier is used, results showed that the performance is reached to 80% when allocating 70% of the images to the training process and 30% of the images to the testing process.

4.2 Future work:

Histogram contributes converting images into a set of features .In the future researches, the following developments can be applied and the results are compared with current system:

1- Extraction of features can be achieved using different method such as moment and GLCM.

2- Using K-means and FCM for clustering phase.

3- Using another tool for classification and such as neural , KNN.

Reference

[1] Kalpana U. Rathod and Y. D. Kapse "Automated Brain Tumor Detection and Brain MRI Classification Using Artificial Neural Network A Review", International Journal of Science and Research (IJSR) ISSN (Online): 2319-7064 , Volume 5 Issue 7, July 2016.

[2] A.Sindhu and S.Meera "An Ann Approach For Classification Of Brain Tumor Using Image Processing Techniques In Mri Images " International Journal Of Engineering Sciences & Research Technology, ISSN: 2277-9655 December,2015.

[3] Hussein A. Lafta and Esraa A. Hussein, " Design a Classification System for Brain Magnetic Resonance Image ", Journal of Babylon University/Pure and Applied Sciences, Vol.21, No.8, 2013.

[4] Rakesh Kumar and Jyotishree , Multiclass Classification of Brain Tumor in MR Images, International Journal of Innovative Research in Computer and Communication Engineering Vol. 4, Issue 5, May 2016.

[5] Sudipta Roy , Shayak Sadhu , Samir Kumar Bandyopadhyay, Debnath Bhattacharyya and Tai-Hoon Kim, Brain Tumor Classification using Adaptive Neuro-Fuzzy Inference System from MRI, International Journal of Bio-Science and Bio-Technology Vol. 8, No.3 (2016), pp. 203-218.

[6] Ch.Amulya and G. Prathibha , MRI Brain Tumour Classification Using SURF and SIFT Features, International Journal for Modern Trends in Science and Technology, Volume: 2 | Issue: 07 | July 2016 | ISSN: 2455-3778.

[7] Noor K. A., Asraa A. H, Zainab F. H." Classification of Brain MRI Images using Classifier Techniques supported by Genetic and Fuzzy C-Means "Research Journal of Applied Sciences Vol.11, No.10, pp.1137-1142,2016.

[8]Naveena H , Shreedhara K , Mohamed Rafi, "Detection and Classification of Brain Tumor using BPN and PNN Artificial Neural Network Algorithms"IJCSMC, Vol. 4, Issue. 4, April 2015, pg.782 – 789.

[9] Prachi Gadpayleand Prof.P.S.Mahajani," Detection and Classification of Brain Tumor in MRI Images",International Journal of Emerging Trends in Electrical and Electronics (IJETEE – ISSN: 2320-9569) Vol. 5, Issue. 1, July-2013.

[10] SivaSankari, S, Sindhu" Feature Extraction of Brain Tumor Using MRI" International Journal of Innovative Research in Science, Engineering and Technology, vol. 3march 2014.

[11] Mandhir Kaur and Rinkesh Mittal. 2014. "an Efficient Scheme for Brain Tumor Detection of MRI Brain Images Using Euclidean Distance With FVT", journal of emerging technologies in web intelligence, vol. 6.

[12] R. J. Deshmukh ,R. S. Khule .2014Brain Tumor Detection Using Artificial Neural Network Fuzzy Inference System (ANFIS), International Journal of Computer Applications Technology and Research, ISSN: 2319–8656, Volume 3– Issue 3.