IJCRT.ORG ISSN : 2320-2882
INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)
An International Open Access, Peer-reviewed, Refereed Journal

Hybrid Approach for Brain Tumour Detection in Image Segmentation: A Review

Sandeep[1], Jyoti Kataria[2]
[1]MTech Scholar, [2]Asstt. Professor
[1,2]Department of Computer Science & Engineering MITM Jevra Hisar(Haryana)

***Abstract*- Analysis of cluster is a descriptive assignment that perceive homogenous group of objects and it is also one of the fundamental analytical method in facts mining. The main idea of this is to present an overview about brain tumour detection system and various data mining methods used in this system. This is review and focuses on scalable big-data systems, which include a set of tools and mechanisms to load, extract, and improve disparate data while leveraging the massively parallel processing power to perform complex transformations and analysis will be measured between gadgets by way of measuring the Euclidean distance among every pair of items.**

Keywords: - Big Data, phrases: Clustering, okay-manner, Euclidean distance, imply

I. INTRODUCTION

Data mining is a powerful method for mining useful patterns or data from image and textual data sets. Medical data mining is very important field as it has significant utility in healthcare domain in the real world. Clustering and Classification are the popular data mining methods used to understand the different features of the health data set. This is focused on understanding different techniques for the detection of brain tumour which is an essential decision making feature and is a part of healthcare application. Brain tumour is a life threatening disease which produces problems like brain damage, loss of memory etc. There exist various data mining techniques for early assessment of brain tumour from scanned brain images.

The emerging Big Data Science term, showing its broader impact on our society and in our business life cycle, has insightful transformed our society and will continue to attract diverse attentions from technical experts and as well as public in general [1] [2]. It is obvious that we are living in Big Data era, shown by the sheer volume of data from a variety of sources and its rising rate of generation. Data mining has become fashionable, not just in computer science (journals & conferences), but particularly in business IT.

The emergence is due to the growth in data warehouses and the realization that this mass of operational data has the potential to be exploited as an extension of Business Intelligence. Data mining is more than just conventional data analysis. It uses traditional analysis tools (like statistics and graphics) plus those associated with artificial intelligence (such as rule induction and neural nets). It is all of these, but different. It is a distinctive approach or attitude to data analysis. The emphasis is not so much on extracting facts, but on generating hypotheses. The aim is more to yield questions rather than answers. Insights gained by data mining can then be verified by conventional analysis.[7]

It's far the computational undertaking to partition a given enter into subsets of equal characteristics. These subsets are generally referred to as clusters. it's miles a primary project of exploratory facts mining, and a common method for statistical information analysis, used in many fields, which includes picture analysis, pattern reputation, system learning, facts retrieval and bioinformatics. Many variations of k-means clustering set of rules have been developed recently. [1]

Data mining combines the concepts, tools and algorithms of machine learning and statistics to analyse very large data sets, so as to gain insight, understanding and effective knowledge and it is applied for this purpose in many companies.

We are working on Matlab tool. However, the main reason for switching to Matlab is to take advantage of the coverage and availability of new applications where Matlab is the best in the areas such as effective generalized mixed models, as well as other general and additional models.

Another reason for learning Matlab may be the desire of users to understand the literature, as more people in various fields of sciences publish their research results in the context of Matlab. This means that everything what is done with Matlab can be saved as an object. Every object has a class. It describes what the object contains and what each function does.

2 BRAIN TUMOUR DETECTION SYSTEM

Brain tumour detection system is one of the health care applications and it is essential for early stage detection of tumour. It is a software based application and it is used for better decision making in health care industry. Brain tumour detection system will make an early diagnosis of the disease based on several methods like data mining, machine learning etc. Most of the existing system consists of training part and testing part for detecting the disease. And it uses scanned brain MRI images as input data and train data. The system may consist of pre-processing stage and diagnosis stage. In pre-processing stage the training and testing MRI images are subjected to various image processing techniques for enhancing their quality. After that this enhanced images are subjected to feature extraction and diagnosis. The diagnosis part is done based on the extracted feature. Such system provides powerful decision making and doctors can use it as a second opinion to detect the disease.

Hybrid Techniques

Several hybrid neuro-fuzzy approaches for MRI brain image analysis are reported in the iterature. A combinational approach of SOM, SVM and fuzzy theory implemented by [Juang *et al.* (2007)] performed superiorly when compared with other segmentation techniques. The SOM is combined with FCM for brain image segmentation [Rajamani *et al*. (2007)] but this technique is not suitable for tumours of varying size and convergence rate is also very low. A hybrid approach such as combination of wavelets and support vector machine (SVM) for classifying the abnormal and the normal images is used by [Chaplot *et al*. (2006)]. This report revealed that the hybrid SVM is better than the kohonen neural networks interms of performance measures. But, the major drawback of this system is the small size of the dataset used for implementation and the classification accuracy results may reduce when the size of the dataset is increased.

A modification of conventional SVM such as Least square SVM for brain tumour recognition is proposed by [Luts *et al*. (2007)]. A Hybrid approach for pattern classification is reported by [Lin *et al*. (2006)]. The combination of SVM and fuzzy rules is experimented in this work. The results revealed that the proposed hybrid approach is accurate, fast and robust.

Fourier series

Fourier series are a powerful tool in applied mathematics; indeed, their importance is twofold since Fourier series are used to represent both periodic real functions aswell as solutions dmitted by linear partial differential equations with assigned initial and boundary conditions. The idea inspiring the introduction of Fourier series is to approximate a regular periodic function, of period T, via a linear superposition of trigonometric functions of the same period T; thus, Fourier polynomials are constructed. They play, in the case of regular periodic real functions, a role analogue to that one of Taylor polynomials when smooth real functions are considered.

The idea of representation of a periodic function via a linear superposition of trigonometric functions finds, according to [1, 2], its seminal origins back in Babylonian mathematics referring to celestial mechanics. Then, the idea was forgotten for centuries; thus, only in the eighteenth century, looking for solutions of the wave equation referring to a string with fixed extreme, introduce sums of trigonometric functions.

However, a systematic study is due to Fourier who is the first to write a 2p-periodic function as the sum of a series of trigonometric functions. Specifically, trigonometric polynomials are introduced as a tool to provide an approximation of a periodic function. Historical notes on the subject are comprised in [3] where the influence of Fourier series, whose introduction forced mathematicians to find an answer to many new questions, is pointed out emphasizing their relevance in the progress of Mathematics. Since the fundamental work by Fourier [4], Fourier series became a very well known and widely used mathematical tool when representation of periodic functions is concerned. The aim of this section is to provide a concise introduction on the subject aiming to summarize those properties of Fourier series which are crucial under the applicative viewpoint.

Indeed, the aim is to provide those notions which are required to apply Fourier series representation of periodic functions throughout the volume when needed. The interested reader is referred to specialized texts on the subject, such as [5–7] to name a few of them. Accordingly, the Fourier theorem is stated with no proof. Conversely, its meaning is illustrated with some examples, and formulae are given to write explicitly the related Fourier series. Finally, Fourier series are shown to be connected to solution of linear partial differential equations when initial boundary value problems are assigned. In the same framework, a two dimensional Fourier series is mentioned.

Discrete Wavelet Transform

We begin by defining the wavelet series expansion of function $f(x) \in L^2(\mathbf{R})$ relative to wavelet $\psi(x)$ and scaling function $\phi(x)$. We can write

$$f(x) = \sum_k c_{j_0}(k)\phi_{j_0,k}(x) + \sum_{j=j_0}^{\infty}\sum_k d_j(k)\psi_{j,k}(x)$$

where j_0 is an arbitrary starting scale and the $c_{j_0}(k)$'s are normally called the approximation or scaling coefficients, the $d_j(k)$'s are called the detail or wavelet coefficients. The expansion coefficients are calculated as

$$c_{j_0}(k) = \left\langle f(x), \tilde{\phi}_{j_0,k}(x) \right\rangle = \int f(x)\tilde{\phi}_{j_0,k}(x)dx$$

$$d_j(k) = \left\langle f(x), \tilde{\psi}_{j,k}(x) \right\rangle = \int f(x)\tilde{\psi}_{j,k}(x)dx$$

If the function being expanded is a sequence of numbers, like samples of a continuous function $f(x)$. The resulting coefficients are called the discrete wavelet transform (DWT) of $f(x)$. Then the series expansion defined in Eqs. and becomes the DWT transform pair

$$W_\phi(j_0,k) = \frac{1}{\sqrt{M}}\sum_{x=0}^{M-1} f(x)\tilde{\phi}_{j_0,k}(x)$$

$$W_\psi(j,k) = \frac{1}{\sqrt{M}}\sum_{x=0}^{M-1} f(x)\tilde{\psi}_{j,k}(x)$$

for $j \geq j_0$ and

$$f(x) = \frac{1}{\sqrt{M}}\sum_k W_\phi(j_0,k)\phi_{j_0,k}(x) + \frac{1}{\sqrt{M}}\sum_{j=j_0}^{\infty}\sum_k W_\psi(j,k)\psi_{j,k}(x)$$

where $f(x)$, $\phi_{j_0,k}(x)$, and $\psi_{j,k}(x)$ are functions of discrete variable x = 0, 1, 2, ... , M – 1.

The Fast Wavelet Transform

The fast wavelet transform (FWT) is a computationally efficient implementation of the discrete wavelet transform (DWT) that exploits the relationship between the coefficients of the DWT at adjacent scales. It also called Mallet's herringbone algorithm. The FWT resembles the two band sub band coding scheme of.

Consider again the multi resolution equation

$$\phi(x) = \sum_n h_\phi(n)\sqrt{2}\phi(2x-n)$$

Scaling x by 2^j, translating it by k, and letting m = 2k + n gives

$$\phi(2^j x - k) = \sum_n h_\phi(n)\sqrt{2}\phi\left(2(2^j x - k\right.$$

$$= \sum_m h_\phi(m-2k)\sqrt{2}\phi\left(2^{j+1}\right.$$

Similarity,

$$\psi(2^j x - k) = \sum_m h_\psi(m-2k)\sqrt{2}\phi(2^{j+1}$$

and can be illustrated with block diagram of Figure 3.

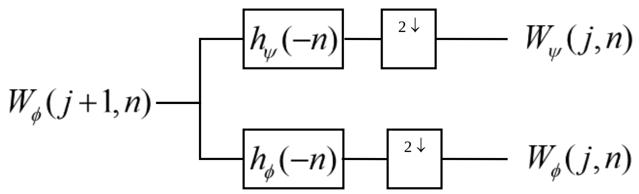

Figure 1 An FWT analysis filter bank.

If function $f(x)$ is sampled above the Nyquist rate, its samples are good approximations of the scaling coefficients and can be used as the starting high-resolution scaling coefficient inputs. Therefore, no wavelet or detail coefficients are needed at the sampling scale.

The inverse fast wavelet transform (FWT^{-1}) uses the level j approximation and detail coefficients, to generate the level j + 1 approximation coefficients. Noting the similarity between the FWT analysis filter bank in Figure 1.1 and the two-band subband analysis portion of Figure 1, the FWT^{-1} have the synthesis filter bank of Figure 1.2.

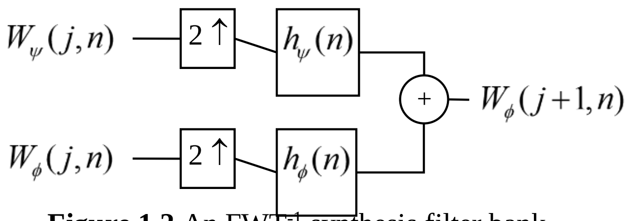

Figure 1.2 An FWT^{-1} synthesis filter bank.

By subband coding theorem of section 1, perfect reconstrucion for two-band orthonormal filters requires $g_i(n) = h_i(-n)$ for i = {0, 1}. That is, the synthesis and analysis filters must be time-reversed versions of one another. Since the FWT analysis filter are $h_0(n) = h_\phi(-n)$ and $h_1(n) = h_\psi(-n)$, the required FWT^{-1} synthesis filters are $g_0(n) = h_0(-n) = h_\phi(n)$ and $g_1(n) = h_1(-n) = h_\psi(n)$.

CONCLUSION

This paper presents an overview of the mean shift and k-means clustering algorithm. K-means clustering is a common way to define classes of jobs within a Dataset. Selection of good initial points will improve solutions and reduce execution time. Therefore the initial starting point selection may have a significant effect on the results of the algorithm. Methods to improve performance of k-means clustering fall into two categories: initial point selection and define number of cluster. We demonstrate that clustering can help infer more coherent topics and can differentiate topics into group-specific ones and group-independent ones.

REFERENCES

[1] Hanna M Wallach. Structured topic models for lan- guage. Unpublished doctoral dissertation, Univ. of Cambridge, 2008.

[2] K. Sravya and S. Vaseem Akram, "Medical Image Segmentation by using the Pillar K-means Algorithm", International Journal of Advanced Engineering Technologies Vol. 1, Issue1; 2013.

[3] Wei Xu and Yihong Gong. Document clustering by concept factorization. In Proceedings of the 27th an-nual international ACM SIGIR conference on Re-search and Development in Information Retrieval,pages 202{209. ACM, 2004.

[4] Wei Xu, Xin Liu, and Yihong Gong. Document clustering based on non-negative matrix factoriza- tion. In Proceedings of the 26th annual international ACM SIGIR conference on Research and Develop- ment in Informaion Retrieval, pages 267{273. ACM, 2003.

[5] Jun Zhu, Li-Jia Li, Li Fei-Fei, and Eric P Xing. Large margin learning of upstream scene understanding models. Advances in Neural Information Process- ing Systems, 24, 2010.

[6] F. Gibou and R. Fedkiw, "A Fast Hybrid k-Means Level Set Algorithm for Segmentation**".** 4th Annual Hawaii International Conference on Statistics and Mathematics 2002, pp. 1-11; 2004.

[7] G. Pradeepini and S. Jyothi, "An improved k-means clustering algorithm with refined initial centroids". Publications of Problems and Application in Engineering Research. Vol 04, Special Issue 01; 2013.

[8] G. Frahling and C. Sohler, "A fast k-means implementation using coresets". International Journal of Computer Geometry and Applications 18(6): pp. 605-625; 2008.

[10]J. Yadav and M. Sharma, "A Review of K-mean Algorithm". International Journal of Engineering Trends and Technology. Volume 4, Issue 7, pp. 2972-2976; 2013.

[11]Sumant V. Joshi and Atul. N. Shire, "A Review of an enhanced algorithm for color Image Segmentation", International Journal of Advanced Research in Computer Science and Software Engineering Volume 3, Issue 3, pp. 435-440; 2013.

[12]J. Wang and X. Su; "An improved K-means Clustering Algorithm". In proceedings of 3rd International Conference on Communication Software and Networks (ICCSN), 2011.

[13]C. Kearney and Andrew J. Patton, "Information gain ranking". Financial Review, 41, pp. 29-48; 2000.

[14]R.V. Singh and M.P.S. Bhatia, "Data Clustering with Modified K-means Algorithm". In proceeding of International Conference on Recent Trends in Information Technology (ICRTIT), 2011.

[15]Li Xinwu, "Research on Text Clustering Algorithm Based on Improved K-means". In proceedings of Computer Design and Applications (ICCDA), Vol. 4, pp.: V4-573 - V4-576; 2010.