International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Kernel Fuzzy C-Means Clustering Algorithm for Detection of Cancer Image Cell using Image Processing Approach

[a]M.Kannan, [b]Nivethitha R*

[a]Assistant Professor, Department of Computer Science, Karpagam Academy of Higher Education
[b]PG Student, Department of Computer Science, Karpagam Academy of Higher Education

A B S T R A C T

Medical imaging often involves the injection of contrast agents and subsequent analysis of tissue enhancement patterns. X-ray angiograms are projections of 3D reality into 2D representations; there is a fair amount of self occlusion among the vessels. Hence one cannot extract the vessels directly using the image intensities or gradients (edge) alone. Vessel extraction from angiogram images is useful for blood vessels measurement and computer visualizations of the coronary artery. This project describes the algorithm for automatic segmentation of coronary arteries in digital X-ray projections (coronary angiograms) .The pattern recognition technique used in this project is K-Means clustering. In this technique clusters are formed based on the minimum distance criteria with random seed point selection. As the dataset's scale increases rapidly, it is difficult to use K-means and deal with massive data, so an Kernal Fuzzy C-Means clustering algorithm is proposed. The performance of the proposed algorithm is compared with other techniques.

Keywords: K-Means, Minimum Distance, Tumor, Segmentation

1. Introduction

A tumor is also known as neoplasm is a growth in the abnormal tissue which can be differentiated from the surrounding tissue by its structure. A tumor may lead to cancer, which is a major leading cause of death and responsible for around 13% of all deaths world-wide. Cancer incidence rate is growing at an alarming rate in the world. Automation of tumor detection is required because there might be a shortage of skilled radiologists at a time of great need. We propose an automatic brain tumor detection and localization framework that can detect and localize brain tumor in magnetic resonance imaging. The proposed brain tumor detection and localization framework comprises four steps: image acquisition, pre-processing, segmentation using k-means clustering algorithm and finally improved k-means clustering technique. After the segmentation operations, tumors appear as pure white color on pure black backgrounds. The proposed tumor detection and localization system was found to be able to accurately detect and localize brain tumor in magnetic resonance imaging. The preliminary results demonstrate how a simple machine learning classifier with a set of simple image-based features can result in high classification accuracy. The preliminary results also demonstrate the efficacy and efficiency of our four-step brain tumor detection and localization approach and motivate us to extend this framework to detect and localize a variety of other types of tumors in other types of medical imagery.

Based on statistics, tumors are the second cause of cancer-related deaths in children (both males and females) whose are under the age of 20 and in males whose age 20 to 39. This disease is also the fifth leading cause of cancer-related deaths in females ages 20-39. This facts increase the importance of the researches on the tumor detection and this will present the opportunity for doctors to help save lives by detecting the disease earlier and perform necessary actions.

Varieties of image processing techniques are available to be applied on various imaging modalities for tumor detection that will detect certain features of the tumors such as the shape, border, calcification and texture. These features will make the detection processes more accurate and easier as there are some standard characteristics of each features for a specific tumor.

All tumors will start small and grow with time. As they grow, they will become more conspicuous and increase the probability of showing their characters. A person with tumor usually has certain symptoms and this will bring that person to a physician. From this, they will be able to detect the smallest possible symptomatic malignant (cancerous) tumors that is in early stage and the smallest possible asymptomatic tumors in the screening process. Basically, there are many factors that can influence the appearance of tumors in different kind of processed images despite some common features of malignancies because of variation in the type of tissue and tumor. For large tumor, characteristic features often to be found while in small tumors, these features of malignancy do not appear to be many and some of them might represent themselves by secondary effects such as distortion in its architecture. In the case of a suspected tumor, “a doctor may perform a neurologic exam to determine if the patient's senses, reflexes, mental status and memory are working normally. Image-based tumor detection uses one or more algorithms as the primary modelling. Some can detect edges; some can detect shapes while others can detect other features. With advances in camera sensing and computational technologies, advances in tumor detection using these features have been an extremely active research area in the intelligent medical community. Clearly, recent researches and trials have extremely help in advancing diagnostic tools for medical purposes but still, the fact that gains in survival need to be achieved by better diagnostic tools.

With the advances in imaging technology, diagnostic imaging has become an indispensable tool in medicine today. X-ray angiography (XRA), magnetic resonance angiography (MRA), magnetic resonance imaging (MRI), computed tomography (CT), and other imaging modalities are heavily used in clinical practice. Such images provide complementary information about the patient. While increased size and volume in medical images required the automation of the diagnosis process, the latest advances in computer technology and reduced costs have made it possible to develop such systems. Blood vessel delineation on medical images forms an essential step in solving several practical applications such as diagnosis of the vessels (e.g. stenos is or malformations) and registration of patient images obtained at different times. Segmentation algorithms form the essence of medical image applications such as radiological diagnostic systems, multimodal image registration, creating anatomical atlases, visualization, and computer-aided surgery

Vessel segmentation algorithms are the key components of automated radiological diagnostic systems. Segmentation methods vary depending on the imaging modality, application domain, method being automatic or semi-automatic, and other specific factors. There is no single segmentation method that can extract vasculature from every medical image modality. While some methods employ pure intensity-based pattern recognition techniques such as thresholding followed by connected component analysis, some other methods apply explicit vessel models to extract the vessel contours. Depending on the image quality and the general image artifacts such as noise, some segmentation methods may require image preprocessing prior to the segmentation algorithm. On the other hand, some methods apply post-processing to overcome the problems arising from over segmentation. Vessel segmentation algorithms and techniques can be divided into six main categories, pattern recognition techniques, model-based approaches, tracking-based approaches, artificial intelligence-based approaches, neural network-based approaches, and miscellaneous tube-like object detection approaches. Pattern recognition techniques are further divided into seven categories, multi-scale approaches, skeleton-based approaches, region growing approaches, ridge-based approaches, differential geometry-based approaches, matching filters approaches, and mathematical morphology schemes. Clustering analysis plays an important role in scientific research and commercial application. This thesis provides a survey of current vessel segmentation methods using clustering approach and provides both early and recent literature related to vessel segmentation algorithms and techniques.

2. Methodology

Original K-means algorithm choose k points as initial clustering centers, different points may obtain different solutions. In order to diminish the sensitivity of initial point choice, we employ a medoid [I1], which is the most centrally located object in a cluster, to obtain better initial centers. The demand of stochastic sampling is naturally bias the sample to nearly represent the original dataset, that is to say, samples drawn from dataset can’t cause distortion and can reflect original data’s distribution In order to lessen the influence of sample on choosing initial starting points, following procedures are employed. First, drawing multiple sub-samples (say J) from original dataset (the size of each sub-sample is not more than the capability of the memory, and the **sum** for the size of J sub-samples is **as** close **as** possible to the size of original dataset) . Second, use K-means for each sub-sample and producing a group of mediods respectively. Finally, comparing J solutions and choosing one group having minimal value of square-error function **as** the refined initial points.

To avoid dividing one big cluster into two or more ones for adopting square-error criterion, we assume the number of clustering is *K'* (K > K, K' depends on the Balance of clustering quality and time). In general, bigger **K'** can expand searching area of solution space, and reduce the situation that there is not any initial value near some extremum. Subsequently, re-clustering the dataset through K-means with the chosen initial conditions would produce K' medicos, then merging K' clusters (which are nearest clusters) until the number **of** clusters reduced to k.

MODULE

There are five modules in our system, they are

- Preprocessing
- Clustering
- Centroid calculation
- Segmentation
- Results

Pre Processing:

The pre-processing involves the resizing, image conversion, image contrast adjustments etc.The image from the MRI cannot be directly used for this operation so some process has to be done to ensure the operation. The first thing has to be done is to resize the image to standard size.

Clustering:

Clustering is a process of collection of object which are similar between them while dissimilar object belongs to other cluster. A cluster technique is used to obtain a portion of N objects using a suitable measure such as resemblance function as a distance measure.

Centroid Calculation:

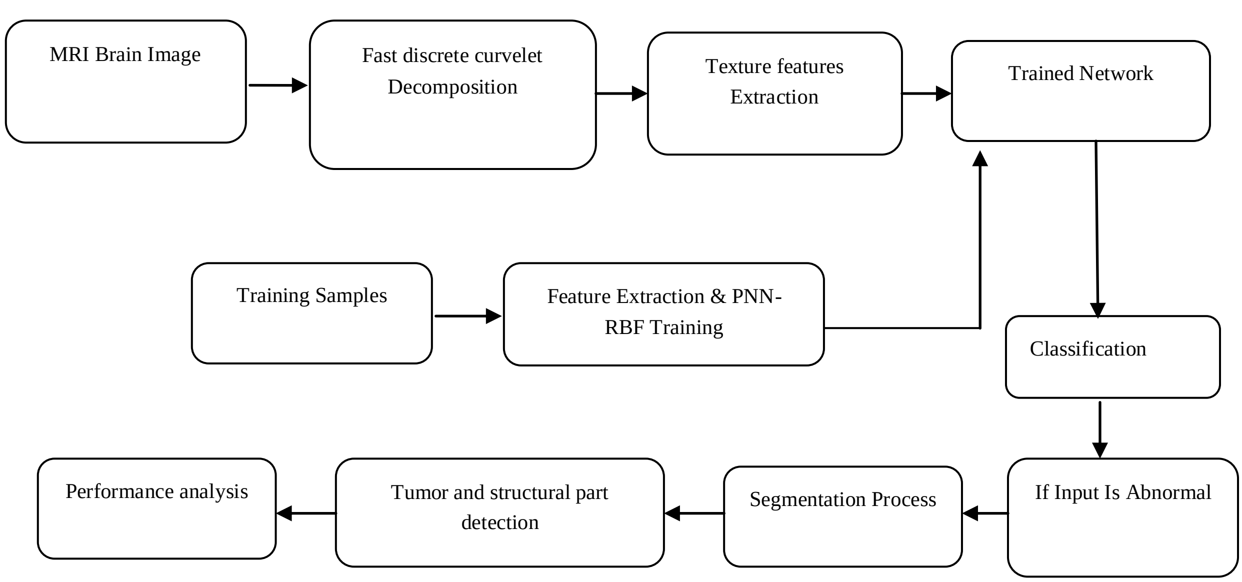
Initial centroid algorithm is useful to avoid the formation of empty clusters. The centroid values are taken with respect to the intensity or mean value of the image.

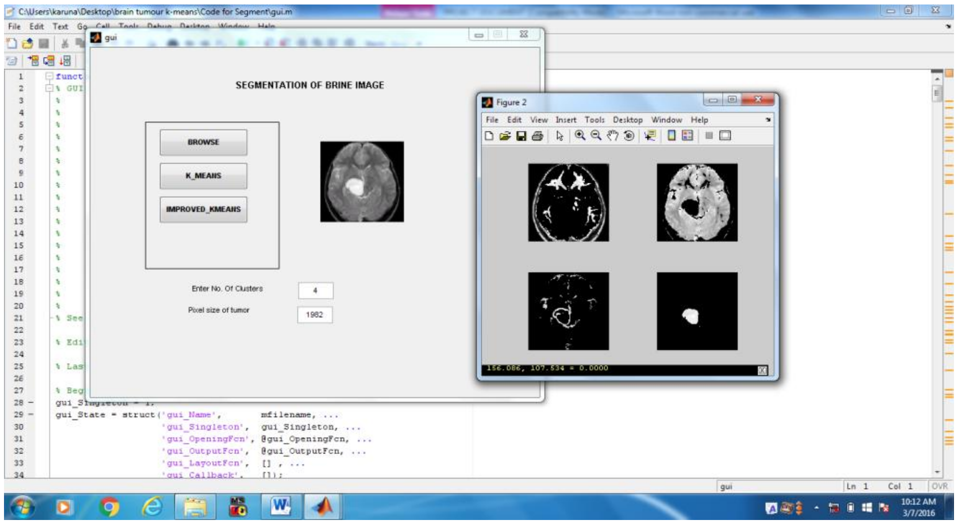
Segmentation:

Segmentation is done using clustering technique, which separates the vessel structure from background. There is no clustering algorithm that can be universally used to solve all problems. So the goal is to search for best algorithms that can be used to segment medical images.

Results:

From the above. It is proved that the tumor is detected and size is found using morphological image processing techniques. The size of the detected tumors can be found using the specifications in the workspace.





Experimental Results

3. Conclusion

Vessel segmentation methods have been a heavily researched area in recent years. Even though many promising techniques and algorithms have been developed, it is still an open area for more research. This algorithm does not require any user interaction, not even to identify a start point. Here seed points are selected randomly which determines the main branches of the vessel structure. Random selection of seed points does not yield accurate segmentation. Accuracy of the segmentation process is essential to achieve more precise and repeatable radiological diagnostic systems. Accuracy can be improved by incorporating a priori information on vessel anatomy and let high level knowledge guide the segmentation algorithm. K-means algorithm is a popular clustering algorithm applied widely, but the standard algorithm which selects k objects randomly from population as initial centroid cannot always give a good and stable clustering. Experimental results show that selecting centroid by our algorithm can lead to a better clustering. Along with the fast development of database and network, the data scale clustering tasks involved in which becomes more and more large. K-means algorithm is a popular partition algorithm in cluster analysis, which has some limitations when there are some restrictions in computing resources and time, especially for huge size dataset. The improved K-means algorithm presented in this paper is a solution to handle large scale data, which can select initial clustering centre purposefully, reduce the sensitivity to isolated point, avoid dissevering big cluster, and overcome defluxion of data in some degree that caused by the disproportion in data partitioning owing to adoption of multi-sampling. Our ongoing research focuses on the development of methods to segment coronary arteries in a sequence of angiographic images while preserving the topology of the vessel structure.

REFERENCES

[1] Cemil Kirbas and Francis K.H. Quek, “A Review of Vessel Extraction Techniques and Algorithms”. Technical report. Vision Interfaces and Systems Laboratory (VISLab), Department of Computer Science and Engineering, Wright State University, Dayton, Ohio, Nov. 2002.

[2] Yu-Fang Zhang; Jia-Li Mao; Zhong-Yang Xiong,” An efficient clustering algorithm” International Conference on Machine Learning and Cybernetics, 2003 Volume 1, 2-5 Nov. 2003 Page(s):261 - 265 Vol.1

[3] Fang Yuan; Zeng-Hui Meng; Hong-Xia Zhang; Chun-Ru Dong; “A new algorithm to get the initial centroids”. Proceedings of 2004 International Conference on**”** Machine Learning and Cybernetics, 2004 Volume 2, 2004 Page(s):1191 - 1193 vol.2

[4] Usama Fayyad, Cory Reina, P. **S.** Bradley Initialization of Iterative Refinement Clustering Algorithms. Microsoft Research Technical Report MSR-TR-98-38, June 1998

[5] Kanungo, T.; Mount, D.M.; Netanyahu, N.S.; Piatko, C.D.; Silverman, R.; Wu, A.Y.” An efficient k-means clustering algorithm: analysis and implementation**”**, IEEE Transactions on Pattern Analysis and Machine Intelligence Volume 24, Issue 7, July 2002 Page(s):881 – 892

[6] D. Guo and P. Richardson, ``Automatic vessel extraction from angiogram images", IEEE Computers in Cardiology, vol. 25, pp. 441-444, 1998.

[7] P.J. Yim, P.L. Choyke, and R.M. Summers, ``Gray-scale skeletonization of small vessels in magnetic resonance angiography'', IEEE Trans. on Med. Img., vol. 19, pp. 568-576, June 2000.
[8] F. Zana and J.C. Klein, ``Robust segmentation of vessels from retinal angiography'', in IEEE International Conference on Digital Signal Processing, vol. 2, pp. 1087-1090, 1997.
[9] R.T. Ritchings and A.C.F. Colchester, ``Detection of abnormalities on carotid angiograms'', Pattern Rec. Let., vol. 4, pp. 367-374, October 1986.
[10] N. Otsu, ``A threshold selection method from gray-level histograms'', IEEE Trans. on Sys., Man, and Cybernetics, vol. 9, pp. 62-66, 1979.
[11] Singh, M.; Patel, P.; Khosla, D.; Kim, T." Segmentation of functional MRI by K-means clustering" Nuclear Science, IEEE Transactions on Volume 43, Issue 3, Part 2, June 1996 Page(s):2030 – 2036
[12] Rui Xu; Wunsch, D "Survey of clustering algorithms" IEEE Transactions on Neural Networks Volume 16, Issue 3, May 2005 Page(s):645 - 678
[13] http://www.elet.polimi.it/upload/matteucc/Clustering/tutorial_html/index.html
[14] http://turing.iimas.unam.mx/~elena/Projects/segmenta/IEEEMI00-R2/node6.html
[15] M Praneesh and Jaya R Kumar. Article: Novel Approach for Color based Comic Image Segmentation for Extraction of Text using Modify Fuzzy Possiblistic C-Means Clustering Algorithm. IJCA Special Issue on Information Processing and Remote Computing IPRC(1):16-18, August 2012. Published by Foundation of Computer Science, New York, USA.