International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-9 Issue-1, November 2019

Brain MRI Segmentation using Cellular Automata in k-Means Algorithm

Jasmeena Tariq, A. Kumaravel

Abstract: ***Tumors in brain have a fast growth(malignant) and should be prevented through various medical ways. However detecting these tumor cells accurately can help the medical professionals to provide accurate and hassle free diagnosis and treatment. Thus we are using Cellular Automata to provide better detection methods in an MRI(Magnetic Resonance Imaging). Cellular Automat is widely used concept with image processing. It is a system which id discreet and dynamic and comprises of simple cellular grid and rules, and works locally. Due to simplicity and usage in complex problems it is widely used concept in many new emerging data science complex problems. Conway's Game of Life is very well known cellular automata and thus researchers are becoming more interested in CA.***

Keywords: Data mining, MRI, Data science, Classifiers, cellular automata, multi-dimensional cellular automata.

I. INTRODUCTION:

This paper uses the concept of Cellular Automata (CA) as a classifier in data mining, using k-means algorithm. Tumor cells within the brain are abnormal cells which show heterogeneous growth (in case of malignant)with the passage of time. They can be distributed in the brain in many small or some big clusters. We have done classification of brain MRI using k-means classifier including cellular automata rules.

Cellular automata has proved to be very useful in self organizing of cells (grids which are usually 1-Dimensional and 2-Dimensional) and can prove beneficial in very complex grids. In a CA all decisions made are done locally, thus having nearly no constraint on modeling and are very simple. Thus CA is considered more useful for data mining. CA is extensively used for pattern recognition and image processing. Parallel multipliers can be built by using CA, as degree of parallelism is very high. Thus simulation of complex systems like data mining becomes easy. With the popularity of ubiquitos computing and high research in nanotechnology, many CA based ideas are used e.g, Quantum-Dot Cellular Automata[6] etc.

Gardner in 1970 introduced, in Scientific American, the best example of CA known as 'Conway's Game of Life'[2].Wolfram et al(1994, 2002) also did some extensive study on elementary Cellular Automata[9][10] . Both of their concepts are widely followed when extending the use of CA in other fields.

CA can be used in the field of mathematics, physics, chemistry, biology and other theoretical fields as well. Many researchers have done their researches with Cellular Automata , like Bandini et al (1996) did research in finding the industrial application of CA, Sullivan & Knight(2004) has applied CA to spreading of forest fire, Syphard Clark & Loulidi(2003) applied CA to urban growth and animal habitats, Benyoussef et al(2003) also applied CA to find the spread of HIV infection in person.CA has also been extensively used for algorithmic ordering problems, image processing , data mining (evolutionary programming and to reveal some genetic diseases). Data mining means the discovery or extraction of useful information in database[3][4][5].We use classifiers to represent data mining. Data mining uses three basic and major classifiers: Neural Networks, if-then rule and decision tree.

II. CELLULAR AUTOMATA

Cellular automata(CA) works dynamically and locally on a grid of cells, represented in matrix form. Gardner in 1970 introduced, in Scientific American, the best example of CA known as 'Conway's Game of Life' [2]. CA is a reversible system working with finite grids and is distributed in nature.

In a CA, each cell should have a finite and discreet value. In CA each cells value is directly dependent on the value of its neighboring cells. All the cells are changed at the same time for each timestamp by following any CA rule. A rule is generally taken from Wolfram's rules (0- 255)[9][10].The various patterns used in elementary Cellular automata is shown in Figure I below:

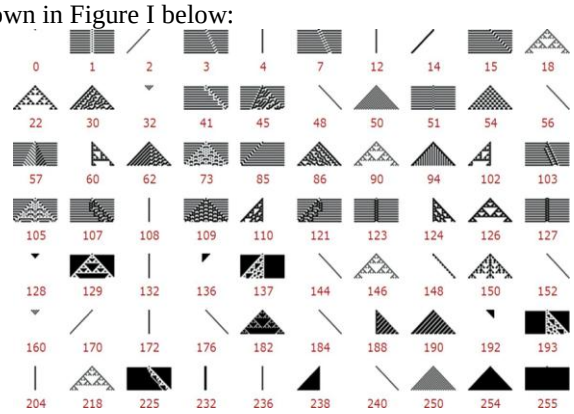


Figure I : Elementary Cellular Automat Patterns.[11]

This CA rule must be same for each cell in a grid. CA is useful in self organization of cells (in a grid, which are usually 1-Dimensional and 2-Dimensional) and has proven quite beneficial in case of complex grids due to its property of locality.

Revised Manuscript Received on November 30, 2019.
* Correspondence Author
Jasmeena Tariq*, Computer Applications, Bharath Institute of Higher Education, Research, Chennai, India.
Dr Kumaravel Professor, Dean, School of Computing, Bharath University, Chennai, India.

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an open access article under the CC-BY-NC-ND license http://creativecommons.org/licenses/by-nc-nd/4.0/

In a cellular automat every cell can have only one finite and discreet value at in a particular timestamp. In a CA all decisions made are done locally, thus having nearly no constraint on modeling and are very simple. We differentiate neighbors on the basis Von-Neuman and Moore neighborhood. Figure I, shows all the three types of neighborhoods in a 2-D Cellular Automata. In this paper common and only simplest CA variations will be discussed, and 2-D grids will be mostly referred.

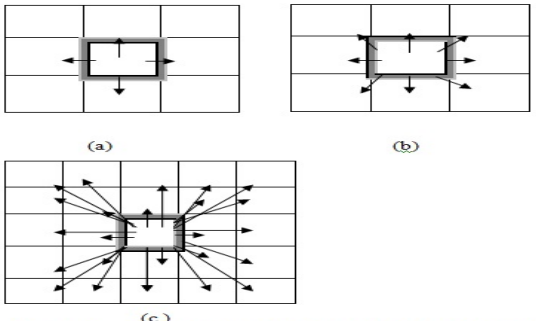


Figure II: a) Von Neuman b) Moore c) Extended Moore Neighborhoods

III. DATA MINING

Data mining means the process of discovering or extracting patterns in data sets where methods of database , statistics and machine learning are involved[3][4][5].Statistics and computer science combinationaly create a subfield with the help of some intelligent methods called data mining.

Data mining is actually more useful in cluster analysis, sequential pattern mining, spatial indices, association rule mining, anomaly detection (unusual records). Data mining can further be used predictive analysis, decision support system and image processing. Data mining does not involve reporting, result interpretation,data preparation and data analysis. Statistical models and machine-learning are used in data mining to uncover hidden patterns in data. In the process of Knowledge Discovery in database(KDD) data mining is an important step. KDD includes selection of data, pre-processing of data, transformation, data mining and lastly interpretation evaluation as its main stages for processing of large data sets.

K-means Algorithm

For clustering k-means algorithm is mostly used for classifying data. It works well when the data is labeled, is noise free and mostly when the data set is small. J. McQueen(1967), developed k-means algorithm. Clustering is efficient image segmentation method in data mining. It is based majorly how the neighborhood of a cluster point(k) is classified.'k' is usually taken as the square root of total number of data points in a dataset and to avoid confusion between two classe of data the value of 'k' is mainly an odd number. It then calculate the distance of k to each data point in the neighborhood. Thus goinf hor re-assigning the data points and then re-calculating values until the desired output is achieved. It provides good results in image segmentation.

IV. PROPOSED METHODOLOGY

In this paper an algorithm is proposed which uses the concept of k-means algorithm and CA rule. Here we have used Conway's Game of Life CA rule. Our first step is to training an algorithm. An MRI is taken as an instance space; this is mapped in a grid, in order to provide the image in cells.It is shown in Figure II below

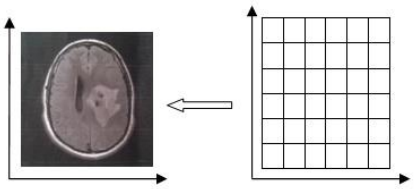


Figure III: Put a grid on instance space

According to game of life we provide the transition rules for the proposed algorithm on this grid:

If value of 'k' at time step 't' is 1 and has maximum neighbors with value 0 at the same time step, then at time step 't+1' value of this cell becomes 0, otherwise it remains 1.

If value of 'k' at time step 't' is 0 and has maximum neighbors with value 1 at the same time step, then at time step 't+1' value of this cell becomes 1, otherwise it remains 0.

In a grid if,

$$a^{t}_{i,j}=1, \text{ then } a^{t+1}_{i,j} = \begin{cases} 1, \text{ if } N_b < \text{Num} \\ 0, \text{ if } N_b >= \text{Num} \end{cases}$$

$$a^{t}_{i,j}=0, \text{ then } a^{t+1}_{i,j} = \begin{cases} 1, \text{ if } N_b >= \text{Num} \\ 0, \text{ if } N_b < \text{Num} \end{cases}$$

Where,

Num=(2^z/2), where 'z' is the size of the grid.

$$N_b = (a_{i,j})^t + (a_{i,j+1})^t + (a_{i+1,j+1})^t + (a_{i+1,j})^t + (a_{i+1j-1})^t + (a_{i,j-1})^t + (a_{i-1,j-1})^t + (a_{i-1,j})^t + (a_{i-1,j+1})^t$$

(for Moore neighborhood, it can be extended in the same way for extended Moore neighborhood)

Process can be shown in Figure III and Figure IV.

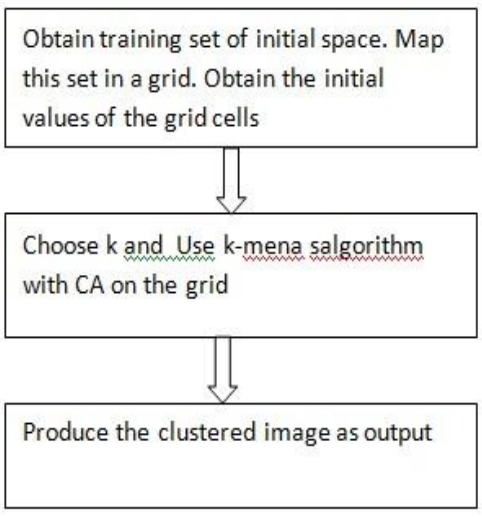


Figure III: Block diagram of our proposed model.

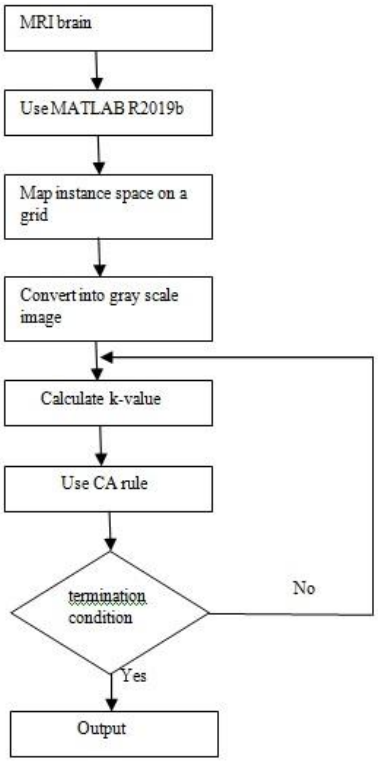


Figure IV: Flowchart of the proposed model

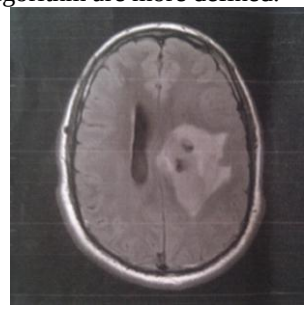
V. ALGORITHM

We have already defined the rules of CA for our algorithm.

1. Input an image in the state space.
2. Divide the whole space containing the image into a grid.
3. Covert it into gray scale image.
4. Determine the initial value of each cell based on its color 0(white) and 1(black)
5. Apply CA rule .
6. For a cluster of cells(e.g 3×3 matrix) in this grid find a cluster cell 'k'.
7. Calculate new cluster point by applying CA rules.
8. If not at the end, re-cluster the data sets.
9. Repeat the steps 7 and 8 until the required output is received.
10. End

VI. RESULT ANALYSIS

The algorithm given above, was run in MATLAB R2019b for both k-means algorithm and our proposed algorithm. The edges in proposed algorithm are more defined.



Original MRI image of brain

Output of k-means algo

Output of proposed algorithm

VII. RELATED WORK

Mayank Arya Chandra et al[6] has also used somewhat same process to show how cellular automat and data mining can be used by elaborating in more detail. In this paper different data mining classifiers with their approach have been discussed briefly. It has shown its main effort on using k-nearest neighborhood algorithm and how cellular automat can be used along with it.

Tungba Usta et al[7] have used heat transfer and state transfer CA as concept in data mining. Their attributes are dependent on neighborhood values.

VIII. CONCLUSION

In this paper CA technique is used with k-means algorithm foe image segmentation of brain tumor MRI. Data science is becoming very important in the field of computer science and in coming age it will gain more importance due to its usage in general problem solving. Thus when cellular automata is blended in data science, a new and easy methods can be applied to problems which are very complex in nature. This paper shows how k-means algorithm can be improved if combined Cellular automat technique is used.includes brief history of data mining and cellular automata.

REFERENCES

1. Tom Fawcett (Center for the Study), "Data mining with cellular automata," pp. 1-10, 2007.
2. M. Gardner. Mathematical games: The fantastic combinations of John Conway's new solitaire game "life". Scientific American, 223:120–123, October 1970. Available: http://hensel.lifepatterns.net/october1970.html.
3. B. Meshkboo and M. Kangavari, "Video Data Mining with Learning Cellular Automata," pp. 1-9.
4. A. Sleit, A. Dalhoum, I. Al-dhamari, and A. Awwad, "Efficient Enhancement on Cellular Automata for Data Mining," Technology, pp. 616-620.

5. L. Zhou and M. Yang, “A Classifier Build Around Cellular Automata for Distributed Data Mining,” 2008 International Conference on Computer Science and Software Engineering, pp. 312-315, 2008.
6. Mayank Arya Chandra, Vidushi, “A study and analysis of Cellular Automata based Classifier in Data Mining”, International Conference on Advances in Computer Applications (ICACA) 2012 Proceedings published by International Journal of Computer Applications.
7. Tugba Usta , Enes Burak Dundar,Emin Erkan Korkmaz,”A cellular automata based classification algorithm”,2018.
8. S. Wolfram. Cellular Automata and Complexity: Collected Papers. Westview Press, 1994.
9. S. Wolfram. A New K ind of Science. Wolfram Media, Inc., Champaign, IL, 2002.
10. http://atlas.wolfram.com/01/01/.

AUTHORS PROFILE

Jasmeena Tariq is a research scholar at, Computer Applications, Bharath Institute of Higher Education and Research, Chennai. Her research interests are Machine Learning, Image Processing, Data mining and Cellular Automata. You can email her at jasmeenat@gmail.com

Dr Kumaravel is working as a Professor and Dean, School of Computing, Bharath University, Chennai. His research interest includes Soft Computing, Cloud Computing, Machine Learning, Pervasive Computing and Knowledge Engineering. He is a life Member of ISTE and IET.