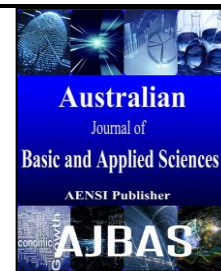




ISSN:1991-8178

Australian Journal of Basic and Applied Sciences

Journal home page: www.ajbasweb.com

Examine and Extraction of Optimized Stem Cells Using Image Processing

¹B. Sreedevi and ²Dr.S.P. Rajagopalan¹Dept. of Computer Science and Engineering/Sri Sai Ram Institute of Technology, Chennai, India.²Dept. of Computer Science and Engineering /GKM College of Engineering and Technology, Chennai, India.

ARTICLE INFO

Article history:

Received 28 January 2015

Accepted 25 February 2015

Available online 6 March 2015

Keywords:

Stem cell algorithm, data clustering, K-means algorithm, Image processing, pattern recognition, mesenchymal stem cell, quality assessment

ABSTRACT

Stem cells have attracted tremendous interest in recent times due to their promise in providing innovative new treatments for a range of diseases. Their ability to regenerate and repair damaged tissue, and hence restore lost body function, in a manner beyond the body's usual healing process. There is presently no quick and easy way to determine the quality of a sample of such cells. There is a great need to assess the quality of cell cultures for research and potential clinical application. The research presented in this uses computerised image processing and an algorithm called Optimized Stem Cells algorithm (OSCA) to provide a quicker and simpler method for the quality assessment. The OSCA aims to obtain a best stem cell that has good self renewal power and minimal cost after n iterations.

© 2015 AENSI Publisher All rights reserved.

To Cite This Article: B. Sreedevi and Dr.S.P. Rajagopalan., Examine and Extraction of Optimized Stem Cells Using Image Processing. *Aust. J. Basic & Appl. Sci.*, 9(10): 1-5, 2015

INTRODUCTION

The selection of stem cells from various organs can be done using data clustering algorithms. Data Clustering is an important issue and a key part of many fields such as data mining, math programming, scientific analysis and image segmentation. Data clustering aims at dividing a dataset into some classes without knowing any pre-information about the kind of relations exists between classes. There are many ways to do data clustering. One of the popular algorithms is the K-means algorithm. The algorithm tries to put the entire dataset S into K clusters (i.e. $C_1, C_2, \dots, C_k$) by randomly selecting K data points as a set of cluster centers. To do that, the clusters are formed such that the existing data in each cluster should have the minimum Euclidean distance to the center of that cluster.

I. Stem Cells Culturing and Imaging:

Adult stem cells are undifferentiated cells that exist in the developed human or animal body. They have the ability to differentiate into the cell types that make up the tissue or organ system from which they originate. As such, they appear to serve as a repair mechanism for the tissue with which they are associated. Adult stem cells are described as multipotent, meaning they have the potential to differentiate into multiple types of cells, although not all the cell types that make up the body. Although more limited in their differentiation potential than embryonic stem cells, adult stem cells may be extracted from a living human donor and thus do not have the same ethical concerns associated with their use.

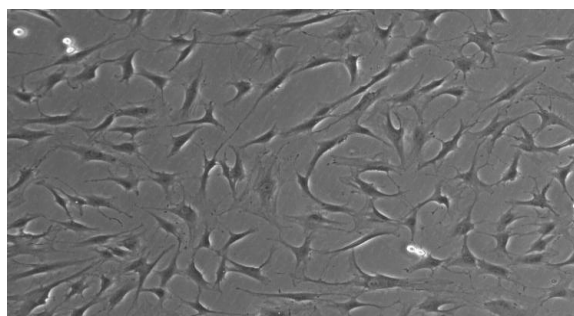


Fig. 1: Human Bone Marrow Stromal Cells-A source of adult stem cells.

Corresponding Author: B. Sreedevi, Dept. of Computer Science and Engineering/Sri Sai Ram Institute of Technology, Chennai, India.
E-mail: sreedevi.balasubramanian@gmail.com

Bone marrow stromal cells (mesenchymal stem cells) are adult stem cells typically found in bone marrow stroma (structural tissue) and other locations throughout the body. They have been shown to be able to differentiate into multiple cell lineages including bone, cartilage, tendon, ligament, fat, muscle (including heart muscle) and neurons. Because of this, they offer hope in the regeneration of tissue damaged by such conditions as

osteoarthritis, Parkinson's disease and for the repair of the heart following heart attack.

Having given a brief introduction to stem cells as well as the current literature relating to their quality assessment, we turn our attention now to the culturing and imaging of stem cells required for use in our project. Equipments used with this paper includes, a microscope (for magnification of cells), camera(for image capturing)

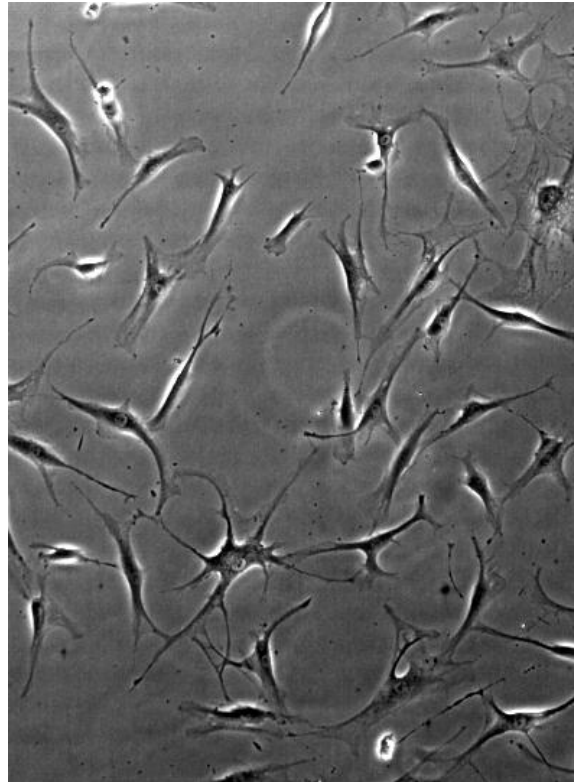


Fig. 2: Sample Image of healthy stem cell captured from Patient-A.

Having now described a validated data , we turn our attention to the design of the image processing and pattern recognition system The first task of the system when given an image is to locate which regions of the image contains cells, and which regions contain background. The CellScan program is a plug-in for the ImageJ application. It is responsible for the image processing (preprocessing, normalisation,segmentation) and feature extraction process of our quality assessment system.

II. Optimized Stem Cells Algorithm:

Most naturally-based optimization algorithms are introduced because of their simplicity and because they have been shown to effectively solve complex optimization problems in nature. Stem cells are found in all multi-cells of body organs where they are able to become a full organ. Research in this field was conducted by Ernest A. Mc Culloch and James E. Till and their activities were focused on investigating embryonic stem cells which are part of the inner blastocysts cells group. In their research they have

reached at an important conclusion which is that stem cells may result in changes in some organs of body. The problems of some adult human organs can also be resolved using the ability of embryonic stem cells in cell division and becoming a full organ.

This algorithm is like other optimization algorithms in that it is based on population and the idea of evolutionary process, but it is different in that it uses minimal constraints and has a simpler implementation than the others. This algorithm converges faster than other optimization algorithms because of its simplicity and its ability to escape from local minima. Population is placed in a range of members (stem cells forming initial population) in this algorithm and it starts with the minimum. Considering that each population member of each stem cell indicates an optimal answer for all considered problems, increasing the population at each iteration is related to the problem space, but defining a large population in this type of algorithms results in abundant iterations to achieve optimal response which consequently raises many problems. Meanwhile defining the population in

an interval and increasing the population according to the space of the considered problem are especially advantageous in implementation by resulting in few iterations in simple problems and increasing the speed of convergence. To perform the proposed algorithm, first, the space of problem is divided into sections which can be done completely in a random process. Then, the initial population is distributed in the whole space of the problem in a random and uniform manner.

In this algorithm, the initial matrix is composed of the stem cell characteristics. It consists of liver cells, intestinal cells, blood cells, neurons, heart muscle cells, pancreatic islets cells and so on.

Thus, the initial matrix is defined as follows:

$$\text{Population} = \begin{bmatrix} X_1 \\ X_2 \\ \dots \\ X_n \end{bmatrix} \quad (1)$$

$X_i = \text{Stem Cells} = [\bar{SC}_1, SC_2, \dots, SC_n]$

$i = 1, 2, 3, \dots, N$

where initialized parameters of the algorithm are defined as:

{
M = Maximum number of stem cells;
P = population size; $\{10 < P \leq M\}$
 SC_{Optimum} = best stem cells in each iteration
 χ = penalty parameter for prevent grow stem cell
 $sc^i = i^{\text{th}}$ stem cell in population
}

This initialization parameter plays a vital role in defining clusters from each part of the body. After determination of the criterion function based on the nature of the problem, the cost of each stem cell is calculated using this criterion function. Two memories are defined for each cell, local memory and global memory. The cost of each stem cell is stored in local memory. A group of cells, representing each area, is defined first. Each cell area, which has the best cost related to other cells, stores the cost in its global memory. This process is also done in other regions. The cost factor is calculated to find out the optimal stem cell from each region, in comparison to its relative power which is defined later. The cost function is as follows:

$$\text{Cost}(SC_i) = \begin{cases} 1/\alpha + f_i & f_i \geq 0 \\ 1 + [f_i] & f_i < 0 \end{cases} \quad (2)$$

where α is positive random number in $[0,1]$ and f_i is the cost value of solution SC_i . For maximization problems the cost can be directly used as a fitness function and cost of each cell is normalized by

$$\text{Cost}_N(SC_n) = \text{Max}[\text{Cost}(SC_i)] - \text{Cost}(SC_n) \quad (3)$$

Then, only the best cells of each area are extracted and employed in a self-renewal process.

The best cell should also have the most power relative related to other cells.

This is obtained from the following equation:

$$P_n = \frac{\text{Cost}_N(SC_n)}{\sum_{i=1}^S \text{Cost}_N(SC_i)} \quad (4)$$

where $\sum_{i=1}^S \text{Cost}_N(SC_n)$ is the total cost of stem

cells. P_n is the comparative power of the n^{th} cell. After calculating comparative power of each cell, its value is saved in memory. Each of these cells shares its information and finally the cells are classified from highest to lowest.

Next, information of the best cells of each area is shared and the cell that has the best cost is selected. This cell has a greater contribution in the self-renewal process of the next iteration than the other cells selected from each area. It should be noted that, however, the entire renewal operation of cells contains just a part of participated population in the next iteration (e.g. 60%) and the rest of the population will be selected randomly using uniform contribution of new cells having no information about considered space. Moreover, cell renewal process depends on the space of the problem, i.e. how big the feature space is. The self-renewal operation of stem cells is computed by:

$$\begin{aligned} SC_{\text{Optimum}}(t+1) &= \zeta(t+1) + SC_{\text{Optimum}}(t) \\ \zeta(t+1) &= \mu(t+1) * \zeta(t) \\ \mu(t+1) &= \begin{cases} \mu(t) * 0.85 & \aleph < 1/5 \\ \mu(t) / 0.85 & \aleph > 1/5 \\ \mu(t) & \aleph = 1/5 \end{cases} \end{aligned} \quad (5)$$

where 't' denotes the iteration number, SC_{Optimum} is the best stem cell in each iteration and $\zeta \in [0,1]$ is a random number, however, it can be considered as fixed value. $\aleph$ is a random number between $[-\mu, \mu]$. If the algorithm cannot improve according to Rechenbergs 1/5 rule, $\mu(t)$ is decreased otherwise $\mu(t)$ is increased to speed up the search.

In the original SCA the distribution of cells is completely random. But in this OSCA the cells are distributed from top to bottom of memory in their priority of cost and self renewal. This helps us to reach an optimized solution before trapping a local minima. In order to create self renewal process more randomly a gamma distribution is used instead of uniform distribution and beta distribution.

Gamma Distribution is (6)

$$g(x; \alpha, \beta) = \frac{\beta^\alpha x^{\alpha-1} e^{-x\beta}}{\Gamma(\alpha)} \quad \text{for } x \geq 0 \text{ and } \alpha, \beta > 0$$

Also, since this algorithm considers the population to be taken in intervals for each iteration, this formula

determines as to which part of the population will be selected for the next iteration.

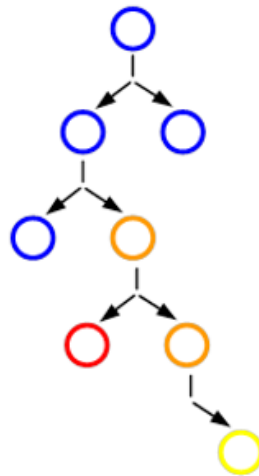


Fig. 1:

Fig.1 shows an example of self-renewal characteristics of stem cells until reaching to the best cell.

This process will continue until it reaches to the goal which is finding the best cell (getting an optimal response) while having the lowest cost function value.

III. Data Clustering Method:

The data clustering algorithms like K-means algorithm, particle swarm optimization (PSO), Ant-colony optimization etc. have drawbacks when factors like population increase and also converge to local minima. The stem cell algorithm is so designed such that, it escapes from any such local minima and the termination is well defined.

After defining required constraints and inserting initial parameters for stem cells algorithm (SCA), the criterion to stop the algorithm should be defined. The algorithm starts with an initial population where each cell in the population indicates a clustering response as K category centers. The initial K centers are selected in a completely random manner without any prior information about the data. Then the distance of each data point to each center is calculated. A data point is considered to be a member of the category with the center to which it has the minimum distance. Then, the difference between each data point (cell) and its center is computed and stored in the local memory of that cell. Note that this difference is a vector. The average of difference vectors in each category is also computed and defined as $\vec{E}_i$. This vector shows how well a center represents its category. The length and direction of this vector quantifies the deviation of the current center from a more representative center. So, we can use this vector to replace the old center by the new and better representative center using the following equation:

$$\vec{Z}_{i-new} = \vec{Z}_{i-old} + \alpha \vec{E}_i \quad (7)$$

where $i \in 1 \dots K$, K is the number of clusters, and α is a learning rate which adjusts the steps in each

iteration towards the better representing center. The learning rate is a positive number with following impact; if it is too small then the convergence might be unnecessarily slow, but if it is too large then the convergence might overshoot and miss the best center. A value between zero and one is recommended. This process will continue until the stopping criterion is met or the number of iterations reaches to the predefined maximum. The pseudo-code for the proposed data clustering algorithm is as follows:

Pseudocode:

Begin

Objective function, $f(x)$, $X = [x_1, x_2, \dots, x_N]^T$

Initialize population, $X_i, \{i=1, \dots, N\}$

Define $\alpha, \beta, \zeta_{min}, \zeta_{max}, \mu, \mathbf{R}$.

Sort the initial population based on the objective function value by equation (2)

Normalize the cost of each cell by equation (3)

Evaluate the comparative power of each cell by equation (4)

While ($t < \text{maximum no. of iterations}$)

Find best stem cell according to fitness value.

Allow best cells to self-renew by equation (5)

Gamma distribution of renewal in problem space by equation (6)

In each iteration: Find best stem cell, replace with respective stem cell and compare with stem cells in previous iteration.

If $f(x_i) < f(x_{optimum})$

Accept and save the new solutions in the memory of stem cell

End if

Search for the new solutions in current iteration and compare with the solutions in previous step

End while

End (OCSA)

The selection of stem cells from various

organs is completely related to mathematical formulae. Also, the effectiveness of the algorithm has been proven mathematically using various other data sets such as Crude oil and Wine. It is to be noted that, this algorithm does not emphasize as to how the stem cells shall grow after being injected into the human body. Also, the growth and response of stem cells in coordination with the human patient is in subject to a doctor's observation.

IV. Conclusion and Future Work:

This paper utilizes the mechanism of selection of the best stem cells from a cluster using the algorithms of data mining. The main goal of this paper is to overcome the shortcomings of the previous clustering algorithms and provide a proper condition for termination of the data clustering process. Also, this algorithm does not utilize the stem cells from any other body, but the patient itself. The stem cell algorithm is efficient as it is not hampered by increasing population nor does it converge to local minima. It effectively arrives at an optimized solution. The methods given above can be implemented in real time.

REFERENCES

- "Stem Cell Basics," National Institutes of Health, U.S. Department of Health and Human Services, April 2009. [Online]. Available: <http://stemcells.nih.gov/info/basics/defaultpage> [Accessed: Mar 23, 2010]
- Abraham, A., S. Das, S. Roy, 2008. Swarm Intelligence Algorithms for Data Clustering, Soft Comput. Knowl. Discov. Data Min., Part IV: 279-313.
- Li, K. and T. Kanade, 2007. "Cell Population Tracking and Lineage Construction Using Multiple-Model Dynamics Filters and Spatiotemporal Optimization," in Proceedings of the 2nd International Workshop on Microscopic Image Analysis with Applications in Biology, Piscataway, NJ.
- Li, K., E.D. Miller, L.E. Weiss, P.G. Campbell and T. Kanade, 2006. "Online Tracking of Migrating and Proliferating Cells Imaged with Phase-Contrast Microscopy," Conference on Computer Vision and Pattern Recognition Workshop, pp: 65-65.
- Li, K., E.D. Miller, M. Chen, T. Kanade, L.E. Weiss and P.G. Campbell, 2008. "Computer vision tracking of stemness," 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, pp: 847-850.
- Merwe Van Der, D.W., A.P. Engelbrecht, 2003. Data clustering using particle swarm optimization, Int. Conf. Evol. Comput., 8-12 Dec., 215-220.
- Nath, S.K., F. Bunyak and K. Palaniappan, 2006. "Robust tracking of migrating cells using four-

color level set segmentation," Lecture notes in computer science, 4179: 920.

Pan, J., T. Kanade and M. Chen, 2009. "Learning to Detect Different Types of Cells under Phase Contrast Microscopy," Microscopic Image Analysis with Applications in Biology (MIAAB).