ANALYSIS OF PERFORMANCE METRICS WITH MESENCHYMAL STEM CELL CLASSIFICATION AND OPTIMIZATION ALGORITHMS

[1] Dr.B.Sreedevi
[1]Associate Professor
[1]Computer Science and Engineering,
[1]Sri Sai Ram Institute of Technology, Chennai, India

Abstract :Stem cells are biological cells found in all multicellular organisms, that can divide (through mitosis) and differentiate into diverse specialized cell types and can self-renew to produce more stem cells. Optimization algorithms have been proved to be good solutions for many practical applications. They were mainly inspired by natural evolutions. However, they are still faced to some problems such as trapping in local minimums, having low speed of convergence, and also having high order of complexity for implementation**.** Mesenchymal Stems or stromal Cells (MSCs) are part of maintaining as well as repairing tissues. The functions are mainly examined in bone marrow-derivedMSC. In the current study, segmentations are performed through usage of Graph-based image segmentations. Feature extraction is performed through Wavelet while Feature Selection is performed through Stem Cell Optimization techniques. Naïve Bayes as well as Support Vector Machines are utilized for classifiers. Results show that the Stem Cell Optimization has better classification accuracy than Information Gain (IG) and Genetic Algorithm (GA).

IndexTerms - Mesenchymal stem cell (MSC), Wavelet, Naïve Bayes (NB), Multi Support Vector Machine (MSVM), Genetic Algorithm (GA), Stem Cell Optimization (SCO).

I. INTRODUCTION

Mesenchymal Stem Cells (MSCs) are multi-potent stem cell populations typically present in bone marrows [1]. The cells inherently differentiate into osteogenic, adipogenic and chondrogenic lineages as well as non-mesodermal lineages which include neural and hepatic lineages. MSCs in adult bone marrow are required for generation of tissues like bones, cartilages, muscles, ligaments, tendons, adiposes as well as bone marrows in the body. Reproducible differentiation of human Embryonic Stem Cells (hESCs) into MSCs requires no usage of feeder layers. MSCs may be grown for sever generations in labs and yet maintain stable morphologies as well as normal chromosome complements.

MSCs refer to non- haematopoeitic stromal cells which are able to differentiate into as well as assist in regenerating mesenchymal tissue such as bones, cartilages, muscles, ligaments, tendons as well as adiposes. MSC is scarce in bone marrows, with around 1 in 10,000 nucleated cells. They are also mortal and so possess the capacity to expand several folds in culture and at the same time retain their growth as well as multi-lineage potentials [2]. MSC may be detected through the expression of several molecules which include CD105 (SH2) CD73 as well as CD34, CD45. The characteristics of MSC ensure that the cell is a perfect candidate for tissue technologies.

In present times, MSC is regarded the most significant tool for therapy in clinics on the basis of its many benefits such as self-regeneration, extended in-vitro expansions, immune-modulation properties, engraftment capacities, multi-lineages differential potentials as well as lesser ethical concerns in comparison to embryonic stem cells. Furthermore, current research suggests that MSC may be isolated from several cell kinds which include adipose tissues, dental pulps, peripheral blood, placentae as well as umbilical cords. The singular characteristics of MSC reveal huge promise in various applications like regenerative medicine, tissues engineering as well as cell-based therapies.

Bone marrow-derived MSCs may be utilized during transplantations for articular resurfacement or supplementation of local stem cell populations of sub-chondral bones. Segmenting refers to the procedure of dividing relevant entities in images from backgrounds. It is required prior to extracting necessary features for usage in the patterns recognition phase. Segmenting is also significant because successful patterns recognition in image processing systems are heavily reliant on the efficacy of earlier segmenting phase.

Basic segmenting was carried out through inputs of image tiles and calculation of standard deviations at all pixels, calculated over circular neighborhoods or masks, centered on pixels. Mask sizes of radius 3, equal to diameters of around 9µm, was discovered in an empirical manner to yield most adequate results [3]. Smaller masks led to more cell areas with lesser intensity variations being wrongly regarded as backgrounds. Larger masks led to more background regions being wrongly sorted into segmented cell region, apart from considerably poorer performance in consequent refined segmenting procedure. Refined segmenting masks yield closer fits to actual cell boundary as opposed to the first rough segmenting contours. Particularly, edge halo regions do not result in false ballooning out of cell boundary contours.

Classification is a common decision making job of human activities. Classificaiton issues occur when objects are to be designated into pre-specfied groups or classes on the basis of the quantity of noted features relating to that object. Several industrial issues are realized as classification issues. The challenge in the resolution of these problems in a mathematical manner rests in the precision as well as distribution of data characteristics as well as model capacities. Patterns recognition is utilized for identifying healthy stem cells utilized for analysis of growth rates. Several techniques for identifying patterns are present. Certain common classification protocols are Support Vector Machines (SVM), Artificial Neural Networks (ANN), Naïve Bayes as well as Decision Tree protocols.

Segmentation of MSC's in the current study is done as Graph Based. The features of the cells like mean, varaiance, standarddeviation, maxamp, minamp, intensity are extracted using Discrete Wavelet transformation. The obtained features are selected with the classification algorithms like Naïve bayes, Support Vector Machine and Stem Cell Optimization algorithms.

II. LITERATURE SURVEY

Sylvester Darvin Sandhaanam .,[4] studied Scientists are now utilizing stem cells of different origin; opening up the research and treatment options for humans. Differing from embryonic stem cells, adult stem cells are procured form a variety of tissues, including skin, fat (adipose) and bone marrow, among other tissues. Adult stem cells are less controversial because the samples are easily obtained and the "host" is not destroyed, as with an embryo.

Stem cells can differentiate into many cell types as they develop, including bone, cartilage, nerves, muscles, and so on. Thus, treatment using stem cells is treatment using stem cells is termed "regenerative medicine" and has many potential uses for a wide variety of diseases and injuries. Stem cells medicine holds much promise for a variety of diseases, including liver, kidney, heart, neurologic and immune-mediated diseases.

Taherdangkoo .,[5] proposed a new optimization algorithm that has overcome the existing problems in the previously introduced optimization algorithms such as genetic algorithm, ant colony algorithm, particle swarm optimization algorithm and artificial bee colony algorithm. The new algorithm is based on natural behavior of stem cells in reproducing themselves. We have tested the proposed algorithm on several Benchmark functions and the obtained results compared with those of applying other optimization algorithms have demonstrated the superior performance of the proposed algorithm.

Boutin& Benoit [6] brought about differentiating stem cells for several regenerative medical applications. Current research indicates the suggested model possesses the capacity for altering gene expressions in hMSCs, and hence provides a way for controlling differentiations in regenerative medical usage.

Safshekan et al., [7] contrasted effects of employing 3 MPa as well as 5 MPa of sporadic hydrostatic pressure on chondrogenic differentiations of human adipose-derived MSC. They were isolated from fat tissue in the human abdomen and studied for expressions of MSC surface proteins as well as multi-lineage differentiations. Real time PCR technique was utilized for examination of expression of collagen II, as a chondrogenicspecific gene. The highest collagen II expression, near that of native cartilage tissues, was achieved in 5 MPa-HP-GF group which suggests that combining loading regimes with utilization of growth factors could be utilized as efficient tools in cartilage tissues engineering for direction of stem cells towards chondrogenic phenotypes.

Choi et al., [2] suggested the examination of single-cell interactions between MSC as well as CD4 T cells innanowell arrays. Utilizing nanowell arrays, it is demonstrated that individual MSC are able to modulate in a dynamic fashion, the T cell proliferation rates, responding to persistent cell-cell interactions in micro-environments. On the basis of the results, the usage of the suggested nanowell array was expected in classifying sub-populations in MSC, resulting in particular therapeutic intervention.

For analyzing impact of Natural Cerebrolycin (NC) on gene expression profiles of MSC, viability of the clinical usage of Alzheimer's Disease(AD) as well as the method of gene molecular Pharmacodynamics were researched by Yinghong et al., [9].Natural Cerebrolycin's anti-AD pharmacodynamicmethod of molecular genes may have relation with several gene expression of certain signal transduction pathway that take part in the regulation of nerve cell differentiation as well as development, learning as well as memory function, and typical cell functions like growth, differentiation, proliferation as well as apoptosis.

Joutsijoki et al., [7] studied the adequacy of hierarchical schemes of multi-class SCMs for classifying induced pluripotent stem cell (iPSC) colony images. iPSC technologies provide immense potential for safe as well as patient-specific drug therapies with

no ethical issues. Two processes for reducing features which are carried out for both data sets are present. During classification, the investigators studied how various hierarchical strategies impact classification. The investigators carried exhaustive evaluations and best precision achieved was 54% with the usage of linear kernel functions. Between various hierarchical structures, in several cases, there are no considerable alterations in outcomes. So, intensity histograms are an excellent base line for classifications of iPSC images, but more nuanced techniques for extracting as well as reducing features along with various other classification techniques are the point for future ersearch.

A new usage of Hidden Markov Models is given by Black et al., [8] who utilized it for helping research meant for testing immunoregulatory impacts of MSCs incynomolgus monkey models of islet transplantations. The model is an unsupervised learning data mining method and it utilized for automated determination of postoperative days (POD) relating to reduction in graft functions, a potential symptom of transplant rejections, on non-human primates after isolated islet cell transplants. Outcomes on expert labeled data revealed the model to be precise 60% of the time. Cluster on the basis of HMM suggests a potential relation between donor haplotype matching as well as loss of function outcome.

III. ARCHITECTURE

The proposed architecture is implemented in MATLAB software and weka tool. The Mesenchymal stem cells (MSC's) are extracted as images from 3 patients A, B and C.The extracted images are added with Gaussian noise so as to enhance the image set to 105 images. The images are segmented into regions using Graph Based Image Segmentation wherein the cells in one region will have lesser weight and cells in different regions seem to have more weights. The Features like intensity, standard deviation, mean, variance, self renewal, comparative power of MSC's are extracted using Wavelets and the required features like comparative power and self renewal were selected using Stem Cell Optimization technique. The images are classified with multi SVM and Naïve Bayes Classifiers. Figure 1 shows the proposed architecture.

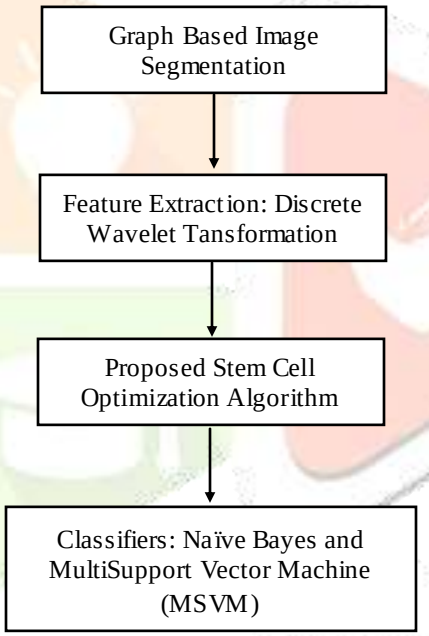

Fig. 1. Architecture

3.1 Segmentation of MSC

In graph-based image segmentation S refers to partitions of V into parts so that all parts or regions $C \in S$ relate to linked parts in graph$G' = (V;E')$, wherein $E' \subseteq E$. edges between two vertices in one part ought to have comparatively less weights, while edges between vertices in two distinct parts are to have greater weights.

Input Matrix M was scanned from top to bottom as well as from left to right testing top as well as left neighbors of all elements of M. If any differed from the current one, they were regarded to denote pixels which were part of separate linked component. The adjacent components must be edges in graphs linking nodes which represent them. Segmentation of MSC's is done with WEKA tool with a sample of 105 images.

3.2 Wavelet based feature extraction

The feature extraction allows extracting list of features that allows discrimination of cells .The list of features extracted from each image is stored as a vector called feature vector. Each feature vector represents the image on the whole. But the selected features are extracted. The intensity based features likemean,median,mode,standard deviation,variance,kurtosis,skew and entropy were extracted from each image.The inner and outer boundaries are identified by taking segmentation mask.

Wavelets refer to waveforms of restricted durations which possess average values of zero. Sinusoids on the other hand, have extensions (in theory) from minus to plus infinities. Wavelets are more suited for description of abnormalities, pulses as well as other events which begin and end within signals. Fundamental idea behind a wavelet is the analysis on the basis of s cale, that

is, analyses are carried out on various scales or resolutions. This allows performance of multi-resolution analysis on image data which reveals texture that might not be distinct else.

Wavelet transforms convolve with a particular signal with a group of shifted as well as scaled variants of original wavelets. At all scales, approximation coefficients are created by low-pass filters as well as detailed coefficient from high-pass filters. Otherwise put, low-pass filtered signals are rough representations whereas high-pass filtered signals comprise details. DWT employed on images are two dimensional transforms which examine images across rows as well as columns which separate horizontal, vertical as well as diagonal details.

3.3 Stem Cells Optimization Algorithm (SCOA)

Stem cell optimization algorithm was inspired by behavior of stem cell function in body. High speed, low level of computation, and simple implementation are the main advantages of proposed algorithm.

This protocol is a population-based EA. It is not an entirely new notion within the domain but improvements in terms of speed of convergence as well as escaping local minima are proposed in this technique. The protocol possesses great flexibility as well as adequate speed in comparison with other optimization protocols, largely. It is to be noted that similar as well as dissimilar proliferation of choosing best stem cells in all iterations as well as its usage is the fundamental core of the protocol. The capacity of stem cells to be attractive is also another significant component of the protocol.

In other protocols, initial population is fixed, whereas in this protocol, population is given in ranges so that protocols can raise population in a gradual manner with every iteration till optimal solution is reached [12]. Taking into consideration the fact that all members of populations are optimized answers for issues, defined populations as well as their rise in every iteration relies on space of specified issue, such that the problem space is split into many regions and stem cells are sent for all regions to explore them. Also taking into consideration that objective of all optimization protocols is the achievement of optimum solution relative to parameters of issue; matrices of parameters are to be generated. Initial matrices of parameters in SCA comprise attributes of stem cells which transfer to organs or tissues of adult individuals.

Initial matrices that are comprised of problem parameters, which are stem cell characteristics which are inherent, are created. For instance, multi-dimensional characteristic may be defined as the capacity of stem cells to transform to marrow cells, blood cells and so on. Initial matrices are given by:

$$SC_i = [SC_{i1}, SC_{i2}, ..., SC_{iD}]$$
$$i = 1, 2, ..., S$$

Wherein S denotes total quantity of cells taking part in implementation procedure of protocol while D refers to problem space dimensions. Initial populations are chosen such that distributions are uniform as well as arbitrarily extended in problem spaces.

Measure function is utilized to find cost of every stem cell. For every stem cell 2 memories are considered: local memory and global memory. Local memory stores cost value of every stem cell. The best cell is the one having the lowest cost for the first iteration. Region wise best cell is chosen. Global memory stores the cost and location of best cells of all regions and from this best cells are selected and utilized in ensuing iterations.

Cost functions for all cells are given by:

$$Cost(SC_i) = \begin{cases} \dfrac{1}{a + f_i} & f_i \geq 0 \\ 1 + |f_i| & f_i > 0 \end{cases}$$

Where in a refers to positive arbitrary number ranging between [0,1], however for other cases with normal complex (such as Dimension (D) <50), a refers to constant with value 1, while fi refers to cost values of solutions SCi. For maximization issues, cost functions may be directly utilized as fitness functions. Costs of all cells are normalized through [12]:

$$Cost_N(SC)_n) = \text{Max}[Cost(SC_i)] - Cost(SC_n)$$

Where in Cost (SCn) refers to costs of nth stem cells, Max [Cost (SCi)] refers to maximal costs amongst stem cell while CostN(SC)n) denotes normalized costs of nth stem cells.

Prominent variable in the determination of best stem cells (optimal solutions) is the relative potency (that is, potential of all stem cells in distinguishing between various cell kinds, for instance, marrow and blood cells). It is got by:

$$P_n = \frac{Cost_N(SC_n)}{\sum_{i=1}^{S} Cost_N(SC_i)}$$

Self-renewal procedure is performed by:

$$SC_{Optimum}(t+1) = \iota \times SC_{Optimum}(t)$$

Where int denotes every iteration while ι refers to an arbitrary number between [0,1].

3.4 Classification of MSC's

Generic method discovered in Bayesian CRM methods like Naïve Bayes that has been vastly employed in machine learning. Typical notion of Naïve Bayes is the usage of knowledge of probability which involves feature values as well as classes in training data sets for producing models of machine learning applications which may then be employed on untested data. The word 'naïve' is utilized for referring to the presumption that conditional probabilities of dataset feature values given classes are not dependent on conditional probabilities of other feature values given the class.

Support Vector Machines (SVM) are typically utilized in patterns recognition as well as objects recognition originally, given sets of points that are part of one of two classes; linear SVMs discover hyperplanes leaving biggest possible fraction of points of same classes on same side, and performing maximization of distance of either class from hyperplanes MSVMs as classification methods has yielded improved classification outcomes when compared to to other typically utilized pattern techniques like maximum likelihood and neural network classifiers. It is typically beneficial for classifying remotely sensed information.

SVM method attempts to discover optimum separating hyperplanes between classes through focus on training samples which are placed at edges of class descriptor. The training samples are known as support vectors. Those other than support vectors are dropped. In this manner, not only are optimum hyperplanes fitted, lesser number of training samples are efficiently utilized; thereby excellent classification accuracy is attained with small training set. The attribute is beneficial particularly for remote sensing data sets as well as for Object based Image Analysis, wherein object instances are typically lesser in number than as in pixel based methods.

IV. EXPERIMENTAL RESULTS

The MSC images from 3 different patients were improved to 105 sample images by adding Gaussian noise with 35 images per patient.

Fig. 1. Sample image - Patient 3

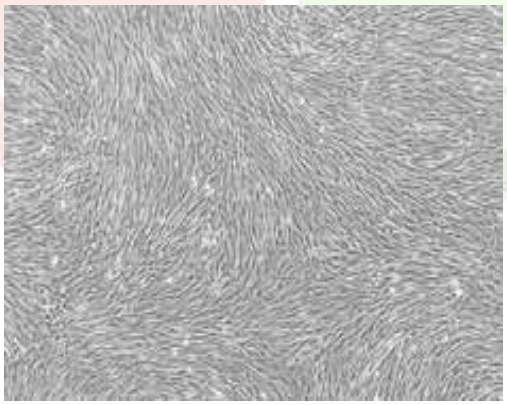

Fig. 2. Masenchymal Stem Cell-Magnified Image

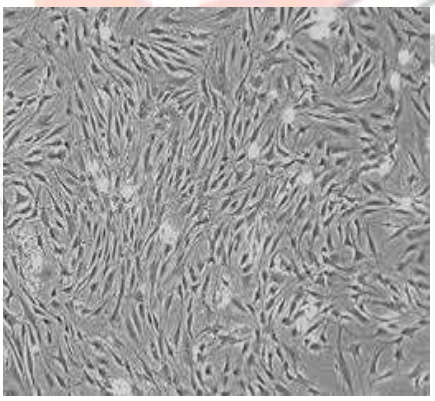

TABLE 1-EXPERIMENTAL RESULTS

Performance Measure	NB-IG	NB-GA	NB-SCOA	MSVM-IG	MSVM-GA	MSVM-SCOA
Classification accuracy (%)	0.86	0.88	0.90	0.91	0.94	0.97
Sensitivity for sample A	0.91	0.91	0.91	0.91	0.97	0.97
Sensitivity for sample B	0.86	0.89	0.91	0.91	0.91	0.97
Sensitivity for sample C	0.8	0.83	0.89	0.91	0.94	0.97

From Table 1, it is observed that the proposed optimization technique with naïve bayes(NB) and MSVM improved the accuracy by 5.5% and 6.0% when compared with IG with naïve bayes and SVM respectively.

V. CONCLUSION AND FUTURE WORK

In this work has been introduced a new optimization algorithm called SCOA, which is based on the behavior of stem cells in reproducing themselves. SCOA was found to have high speed of convergence, low level of complexity with easy implementation process. It could also avoid the local minima in an intelligent manner. MSC is currently used for the treatment

of a variety of clinical conditions. Segmentation of images is a challenging issue. Successful segmentation is an important precursor to successful pattern recognition. A novel approach was proposed in this study for optimized cell image segmentation. For experiments, stem cell optimization was proposed with classifiers such as naïve Bayes and MSVM. The focus of the present work is on the optimization algorithm with MSCs. The Future work can be carried out with the stem cells obtained from adult's different organs like brain, peripheral blood, blood vessels, skeletal muscle, skin, teeth, heart, gut, liver, ovarian epithelium, and testis. The input data set can be tested with other feature selection algorithms. Also we can employ classifiers like ANN and so on.

.

REFERENCES

[1] Prasajak, P., &Leeanansaksiri, W. (2014). Mesenchymal stem cells: current clinical applications and therapeutic potential in liver diseases. Journal of Bone Marrow Research, Volume 2, Issue 1, pp: 1-9.

[2]Choi, J., Hwang, M. P., Lee, J. W., & Lee, K. H. (2012, November). Interactions between mesenchymal stem cells and T cells on a single cell level a nanowell array. In Nano/Molecular Medicine and Engineering (NANOMED), 2012 IEEE 6th International Conference on (pp. 111-116). IEEE.

[3] Iskandar, M. E., Cipriano, A. F., Lock, J., Gott, S. C., Rao, M. P., & Liu, H. (2012, August). Improved bone marrow stromal cell adhesion on micropatterned Titanium surfaces. In Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE (pp. 5666-5669). IEEE.

[4] Sylvestar Darvin Sandhaanam, Ganesan Pathalam, Sudarsanam Dorairaj, Vincent Savariar
Mesenchymal stem cells (MSC): Identification, Proliferation and Differentiation – A Review Article
P.G & Research Department of Advanced Zoology and Biotechnology. Loyola College, Chennai, Tamil Nadu, India.

[5] Taherdangkoo, M., Yazdi, M., & Bagheri, M. H. (2011). Stem cells optimization algorithm. In Bio-Inspired Computing and Applications (pp. 394-403).

[6] Boutin, M. E., & Benoit, D. S. W. (2011, April). A polymeric RNAi delivery system to induce differentiation in mesenchymal stem cells. In Bioengineering Conference (NEBEC), 2011 IEEE 37th Annual Northeast (pp. 1-2). IEEE.

[7] Joutsijoki, H., Rasku, J., Haponen, M., Baldin, I., Gizatdinova, Y., Paci, M., ...&Iltanen, K. (2014, December). Classification of iPSC colony images using hierarchical strategies with support vector machines. In Computational Intelligence and Data Mining (CIDM), 2014 IEEE Symposium on (pp. 86-92). IEEE.

[8] Black, E. F., Marini, L., Vaidya, A., Berman, D., Willman, M., Salomon, D., ...& McHenry, K. (2014, October). Using Hidden Markov Models to Determine Changes in Subject Data over Time, Studying the Immunoregulatory Effect of Mesenchymal Stem Cells. In e-Science (e-Science), 2014 IEEE 10th International Conference on (Vol. 1, pp. 83-91). IEEE.

[9] Yinghong, L., Zhengzhi, W., Ming, L., Xiuqin, J., Min, Y., &Manyin, C. (2009, October). Analysis of Effects of Natural Cerebrolysin on the Gene Expression Profile of Mesenchymal Stem Cells in Rats by Genome Chip and Bioinformatics. In Image and Signal Processing, 2009. CISP'09. 2nd International Congress on (pp. 1-9). IEEE.

[10] Safshekan, F., Hemmati, A., Tafazzoli-Shadpour, M., Shokrgozar, M. A., Haghighipour, N., &Mahdian, R. (2012, December). Effect of hydrostatic pressure amplitude on chondrogenic differentiation of human adipose derived mesenchymal stem cells. In Biomedical Engineering (ICBME), 2012 19th Iranian Conference of (pp. 131-134). IEEE.